Exploring Best-Fit Hazard Functions and Lifetime Regression Models for Urban Weekend Activities: A Case Study

Ming Zhong[1] and John Douglas Hunt[2]

Abstract: Activity-based travel demand forecasting consists of modeling activity type, location, and duration with a view to improving transportation planning and creating effective traffic management systems. Research to date has focused primarily on weekday activity patterns, but given its steady increase, weekend activities and related travel demand also deserve attention. Limited research studied weekend activities, and none of them was found to provide detailed specifications with respect to best-fit hazard functions and lifetime regression models. This study, which took place in Calgary, Alberta (a Canadian city of 1,000,000+), is meant to address that gap. Ten activity patterns of eight demographic groups were assessed and nearly 13,000 observations analyzed. Results affirm that most weekend activities are neither work nor school related and tend to begin mid-day or later; analysis of activity participation by demographic group shows that adults (19–64 years old) are the most active components of our society. Likelihood ratio tests confirm that a two-level modeling exercise is required to handle the heterogeneity within the data: first, analysis by activity type and second, analysis by demographic group. Eleven candidate hazard functions were examined for 10 weekend activities such as shopping or entertainment, then best-fit hazard functions and lifetime regression models were determined. The results show a high degree of fit. It was found that the best-fit parametric models for demographic subgroups are generally consistent with those based on activity type at the aggregate level, a discovery that should simplify future applications. Lifetime regression models show that the starting time of a given activity and personal mobility are the most significant factors influencing activity duration. The applicability of fully parametric, nonparametric, and semiparametric model is discussed and addressed at various points within the paper. The rounding problem of reported durations is also noticed and discussed during the process of identifying best-fit hazard functions and lifetime regression models.

DOI: XXXX

CE Database subject headings: Travel demand; Travel patterns; Transportation models; Regression models; Urban areas; Case studies.

Author keywords: Travel demand; Travel patterns; Transportation models; Regression models.

Introduction

Travel demand modeling has focused on forecasting individual activities and related travel patterns with the purpose of dampening traffic congestion at peak early morning and evening weekday hours. Given increasing travel demand and limited space for construction of new infrastructure, however, congestion has become problematic at or near recreational areas, major shopping centers, sports arenas, and bridges on weekends, especially in large cities [Federal Highway Administration (FHWA) 2004a]. Review of the literature shows that although many have investigated weekday activities (Doherty et al. 2002; Gärling et al. 1994; Miller and Roorda 2003), relatively few have explored weekend patterns (Allison et al. 2005; Bhat and Lockwood 2004; Bhat and Srinivasan 2005; Sall et al. 2005). Those that have done so [Federal Highway Administration (FHWA) 2004b; Parsons Brinckerhoff (PB) Quade and Douglas, Inc. 2000] found weekend household travel, in terms of trip length and number, to be comparable to that of weekdays. Travel behavior, however, differs on weekends as many more are able to participate in personal and social activities. Obviously, further attention to weekend travel is warranted.

A large survey of weekend household activity was completed in 2001 in Calgary, Canada (Hunt and Atkins 2004) with the purpose of collecting data for short-term traffic analysis and long-term transportation planning. The comprehensive data obtained, which permits analysis of weekend versus weekday activities and related travel behaviors, provide helpful insights for policymaking.

In this study, a literature review of duration/hazard models and their applications for modeling household activities, especially weekend activities, is first provided. Next, a brief introduction to study data and a preliminary analysis of weekend activity pattern and participation rate are presented. A large number of hazard functions (HFs) and lifetime regression models are then examined for each activity and demographic group, and best-fit models are specified. Finally, discussion of model options, major findings and conclusions, and future research are given.

[1]Associate Professor, Dept. of Civil Engineering, Univ. of New Brunswick, 17 Dineen Dr., P.O. Box 4400, Fredericton, NB, Canada E3B 5A3 (corresponding author). E-mail: ming@unb.ca

[2]Professor, Dept. of Civil Engineering, Univ. of Calgary, 2500 University Dr., N.W., Calgary, AB, Canada T2N 1N4. E-mail: jdhunt@ucalgary.ca

Note. This manuscript was submitted on March 15, 2007; approved on October 26, 2009; published online on XXXX XX, XXXX. Discussion period open until August 1, 2010; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering***, Vol. 136, No. 3, March 1, 2010. ©ASCE, ISSN 0733-947X/2010/3-1–XXXX/$25.00.

Literature Review

Review of Parametric/Nonparametric/Semiparametric Duration Models

Statistical analysis of lifetime, survival time, and failure time data—known as duration or hazard modeling—has long been crucial to biomedical engineering, the social sciences, and other disciplines. Duration models are being increasingly used in transportation studies (Bhat 2000; Hensher and Mannering 1994).

Both accelerated-time and proportional-hazard parametric models have been used to analyze lifetime data (Bhat 2000; Lawless 2003). The accelerated-time models assume that the effect of covariates is equivalent to altering the rate at which time passes, whereas the proportional-hazard models assume the covariates affect the HF for T (Lawless 2003). Among these models, the following have demonstrated their usefulness and been applied in a wide range of situations (Lawless 2003): exponential, Weibull, loglogistic, lognormal, extreme value (Gumbel), and gamma. For a more thorough discussion of parametric duration models, see Lawless (2003).

The aforementioned HFs are fully parametric and can be applied to questions with a sound theoretical foundation. If little or no knowledge of the functional form of the hazard is available, a nonparametric approach—for which there are no assumptions concerning distribution of the baseline hazard—may be tried. It must be noted, however, that the analyst cannot incorporate explanatory variables for policy analysis with such an approach. The most popular nonparametric method is the Kaplan-Meier estimator. For such an approach, the duration scale is split into small discrete periods; by assuming a constant hazard within each period, the continuous-time step function hazard shape may be estimated. The Kaplan-Meier estimator is particularly good in situations where we want to compare a small number of groups to check if they have similar survival distributions. Several statistical methods, such as the logrank and Wilcoxon, exist for such tests. Also, the nonparametric shape obtained from the Kaplan-Meier estimator is particularly useful for empirically testing assumed parametric baseline shapes (Lawless 2003).

When there is no clear choice concerning HFs, the use of semiparametric hazard models is a relatively safe approach (Lawless 2003). Although the distributional assumption for the baseline hazard may be arbitrary, certain assumptions may be made about the functional form: how, for example, external covariates interact with the model's baseline hazard. Two parametric forms, the proportional and the accelerated lifetime, are usually employed to accommodate the effect of external covariates on the hazard (Lawless 2003). The best-known semiparametric proportional-hazard model was introduced by Cox (1972). Semiparametric models allow for relaxation of the assumption of a parametric relationship between the various factors and the resulting hazard rate. It should be noted, however, that the multiplicative form of interaction between the baseline hazard and external covariates is a strong assumption that requires careful checking when applied (Lawless 2003). If the hazard is generated from a known distribution when a semiparametric model is employed, statistical efficiency will be lost. Whenever information about hazard distribution is not used, a higher error rate and less precise coefficient estimates result (Hensher and Mannering 1994).

Duration Modeling of Activity Patterns

Using data from the San Francisco Bay Area Travel Survey of 2000, researchers at the University of Texas at Austin first drew attention to weekend travel activity patterns. Lockwood et al. (2005) analyzed average frequency and duration, time of day, mode of travel by trip purpose, trip distance by purpose, volume of travel by trip purpose, sequence of activity episodes, activity episode chaining, and purpose of the first and last daily out-of-home episodes. Although this study presented a relatively comprehensive picture of weekend activities in the Bay area by means of statistical analyses, neither duration models nor assessment of lifetime regression for the activities was provided.

Sall et al. (2005) and Bhat and Srinivasan (2005) proposed a weekend activity analysis framework, separating their analyses into the following highest, medium, and lowest levels: pattern, tour, and episode. At the pattern level, the number of daily nonwork/nonschool stops was estimated by means of a multivariate, ordered, and response choice model. The sequencing of all activity episodes and the number of in-home episodes were then examined using multinomial logit models. At the tour level, mode choice was the only attribute considered; a discrete choice framework was used. At the episode level, the researchers employed hazard-based duration models to determine the first (morning) home-stay duration; modeled travel time according to episode and episode duration using simultaneous linear regression; and located stops by means of a disaggregate spatial destination choice model. Although methodologies for duration modeling were proposed, no statistical results were given. Recent work of the Austin groups focuses on modeling frequency of participation in weekend activities (Bhat and Lockwood 2004; Bhat and Srinivasan 2005).

Researchers have increasingly applied duration models to assess nonwork activities. Ettema et al. (1995) used competing risk hazard models to model activity choice, timing, sequencing, and duration of activities of 39 students at Holland's Eindhoven University of Technology. Using New York household survey data from 1997/1998, Chu (2005) employed a Type II Tobit model to analyze nonwork activity durations of workers. Hamed and Mannering (1993) estimated work-to-home travel time with ordinary and three-stage least-squares regression models, reporting a corrected R^2 of 0.11 for ordinary and 0.188 for the latter. When they estimated the Seattle commuter work-to-home departure delay time, Mannering and Hamed (1990) advised using a duration model based on the Weibull distribution. The choice of the Weibull distribution is based on their finding that the end of a departure delay can be viewed as being induced by any one of a number of random factors, such as decreased homeward-bound congestion, boredom with the activity undertaken, and early completion of the activity. Because end times of departure delays depend on the shortest time to the occurrence of one of these random factors, they argued, it should follow the distribution of the smallest extreme. The Weibull distribution method is therefore appropriate. They achieved a standard error of 0.148 for their duration parameter estimates.

Hamed and Easa (1998) developed a Weibull-based proportional-hazard model for modeling urban shopping durations in Amman, Jordan within a larger integrated modeling framework. They reported the significance of the following factors in influencing shopping activity durations: the presence of children, transportation mode, household income, commuter's age and gender, origin of shopping, distance to shopping destination (travel time to shopping), postactivity type, and time of day and

Table 1. Descriptive Statistics for Activity Durations Based on Activity Type

	N	Mean	SDV	COV (%)	Minimum	Median	Maximum
Travel related, etc.	1,019	6.218	14.094	226.67	1	2	311
Work	1,252	241.31	158.32	65.61	1	225	990
School	182	191.84	113.15	58.98	2	180	610
Shopping, etc.	4,539	42.923	45.404	105.78	1	30	780
Sociality, etc.	1,474	137.69	129.89	94.33	1	95	1,440
Eating	1,604	56.96	43.58	76.51	1	45	390
Entertainment, etc.	1,603	135.41	97.62	72.09	1	120	660
Exercise	407	99.7	59.4	59.58	1	90	430
Religious, etc.	787	124.09	75.08	60.5	1	105	600
Out-of-town	15	580.6	335	57.7	60	690	1,110

number of people in the vehicle. They found that the heterogeneity within the data was significant and females tended to spend longer time than males in shopping. They justified their choice of the Weibull-based hazard model with its attractive property of being able to model a monotonically falling or increasing risk. They found that the shape parameter of the Weibull duration model was positive and greater than 1 and thus indicating an increasing HF (Hamed and Easa 1998).

There have been numerous other applications of duration modeling in the transportation area. Nam and Mannering (2000) used hazard-based duration models to evaluate the time of incident detection/reporting, response, and clearance in Washington State; Wang et al. (2002) used fuzzy logic based Weibull duration models for studying vehicle breakdown time on motorways in the U.K. The fuzzy logic approach is to account for subjective, ambiguous, and uncertain information presented in the accident reporting system. Paselk and Mannering (1994) used loglogistic duration models to study vehicular delay at U.S./Canada border crossings. Stathopoulos and Karlaftis (2002) examined the four most widely used HFs for modeling congestion durations in Athens, Greece, and found that the loglogistic form is most appropriate. Gilbert (1992) employed a Weibull duration model to study length of car ownership. Hensher and Mannering (1994) provided a comprehensive review of applications of hazard-based duration models in transport analysis; Lawless (2003) reviewed applications in areas other than transportation. For further details, please see Lawless (2003).

Data and Primary Analyses

Data set from 2001 Calgary Household Activity Survey is used in this study (Hunt et al. 2005). The data used in this study includes personal type (demographic group), employment status (full- or part-time), annual income level, gender, age, household size (the number of people in the household), driving capability (holding driving license or not), activity type, activity duration (in minutes), and start/end times of activity. The investigation comprised 10 types of activities, including (1) travel-related activities such as commuting, drop off, or pick up; (2) work; (3) school; (4) shopping; (5) sociality (getting together with friends or family); (6) eating; (7) entertainment; (8) exercise; (9) religious and civic activities; and (10) out-of-town travel. There are eight personal types or demographic subgroups defined in the data, which are primarily based on people's age and their socioeconomic status. These demographic subgroups are: adult nonworking (AO), adult worker needing car (AWNC), adult worker, no need of car (AWNNC), K-9 students (KEJS), 10–12 students (SHS), postsecondary students (PSSs), seniors 65+ (SEN), and young other (YO). The 10 annual household income categories used were: less than \$25,000; \$25,000–\$35,000; \$35,000–\$45,000; \$45,000–\$55,000; \$55,000–\$65,000; \$65,000–\$75,000; \$75,000–\$100,000; \$100,000–\$125,000; \$125,000–\$150,000; and more than \$150,000. These activity types and identifiers, demographic groups, and annual income levels were defined in the Calgary survey and used here directly. There were 12,916 observations; once those with missing durations were excluded, 12,882 remained for analysis.

One of the first issues to be addressed was how to classify data to model activity durations. Three approaches exist and they are: by individual activity, by demographic group, or by both. Various statistical and plotting methods were employed to check at which level the modeling should be done. Box plots were used to view the distribution of activity duration based on activity and personal type (the results are not shown here). Substantial differences were found between mean levels of activity duration; whereas a travel-related activity (dropping someone off, for example) might take a few minutes, a work period typically lasts 200 min or more. Table 1 shows the summary statistics for the 10 activities studied. Large differences between mean and median durations for different activities suggest such activities should be modeled separately. A likelihood ratio test (LRT) was also carried out to test if a full model including all activities is transferable, that is, to test whether a single model can be used to model durations of all types of activities together (Washington et al. 2003). Study results show that a full model has a much smaller loglikelihood (−49,406) than the sum of all loglikelihood of the submodels (−45,067) developed for each individual activity. The critical chi-square value at 95% confidence level with a degree of freedom (DOF) of 163 is $\chi^2_{0.95}(163)=193.8$, whereas the test statistic is calculated as $\lambda=-2(-49,406+45,067)=8,676$. Since $\lambda \gg \chi^2_{0.95}(163)$, the hypothesis of a full model with transferability is rejected.

A Cochran's Q test (Gavaghan et al. 2000; Higgins et al. 2003) was also used to check if there is significant heterogeneity within the data (results are not shown here). The p values of inlying and outlying variance test are found all less than 2.2×10^{-16} and thus further confirm the presence of heterogeneity. Further analysis for the equal shape parameters across activity groups resulted in a likelihood ratio chi-square statistic of 1,080.36 with a p value of 0.000. All of the aforementioned statistics support that the preceding activities should be modeled separately. Another discovery from the data was that activity duration distributions skew to the right, which indicates models based on distributions other than normal are desirable (Weibull, for example).

Figs. 1(a and b) show weekend and weekday household activ-

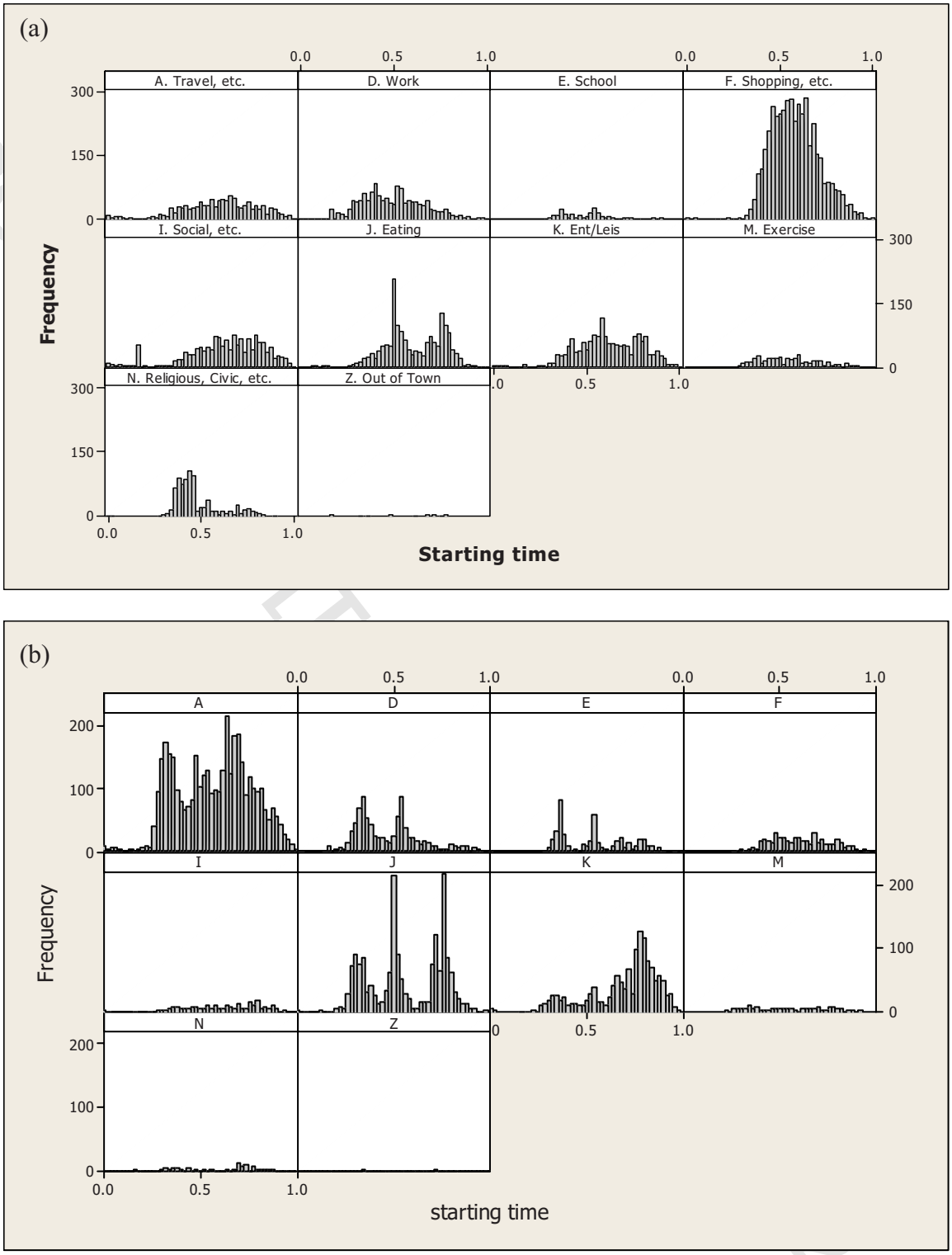

Fig. 1. Comparison of household: (a) weekend; (b) weekday activity patterns

ity patterns. The magnitude of activities are consistent with expected weekly periodicity: Fig. 1(a) shows a much stronger participation of "typical" weekend activities, such as shopping, sociality, and religious, civic, etc.; whereas Fig. 1(b) shows that a higher level of the typical weekday activities, such as travel related (commuting, pick up or drop by), work, school, and out-of-town travel (for business purpose). Please note that a larger scale is used for presenting the weekend activities. With the exception of religious and civic activities, Fig. 1(a) reveals that most weekend activities begin in the afternoon (i.e., after 0.5 in the figure which represents the noon). It is also interesting to notice that weekend activities tend to have fewer peaks than their weekday counterparts. For example, there are three peaks for the weekday commuting travel (A) corresponding to morning, noon, and evening peaking hours, but there is only one very small peak for the weekend travel. Similar finding can be observed for the school

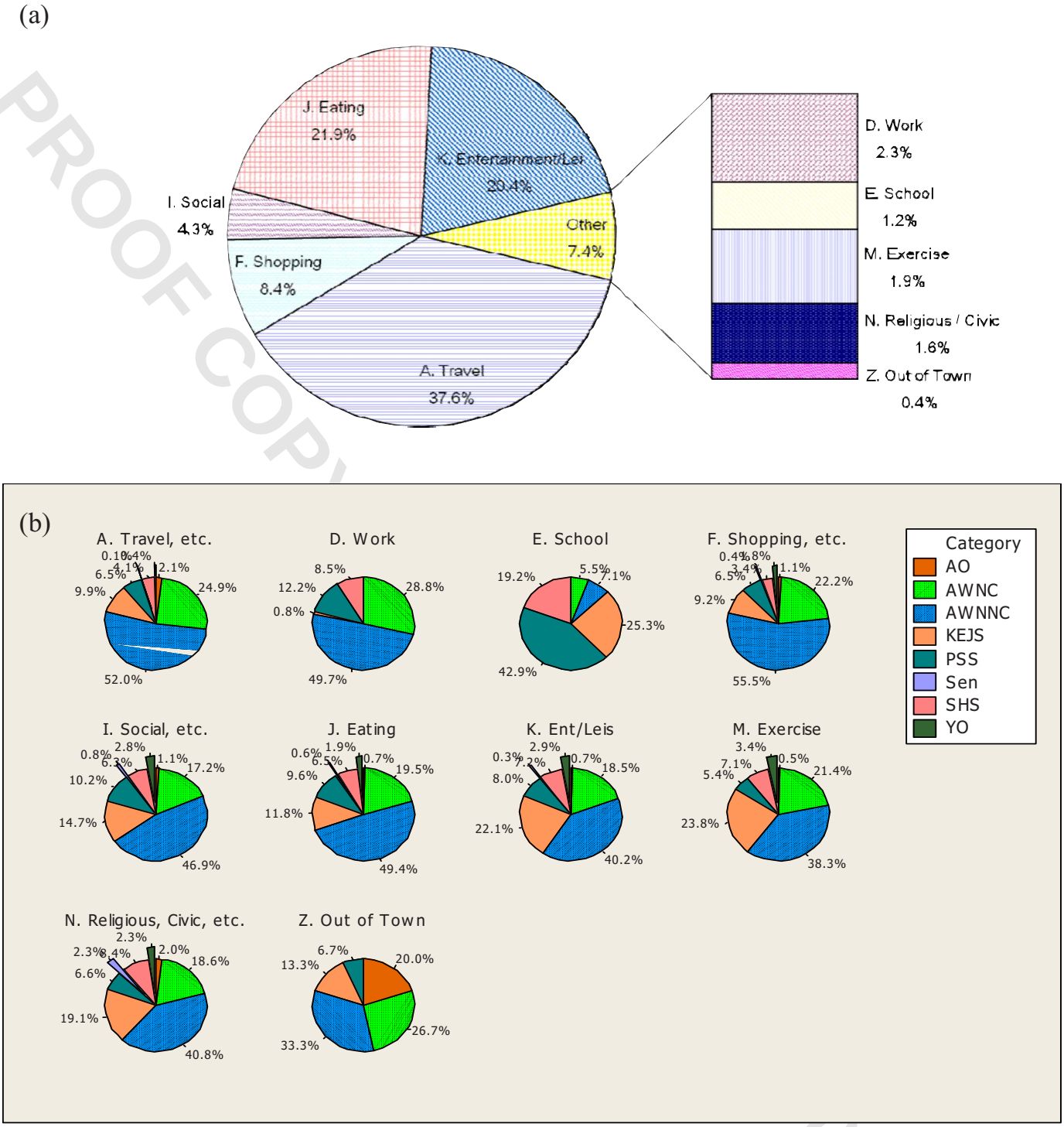

Fig. 2. (a) Percentage of different weekend activities; (b) demographic participation

(E) and out of town (Z) activity, as people tend to start their out-home activities late over the weekends. As is shown in Fig. 1(a), intensity of weekend activity over a short period in the afternoon indicates a relatively high traffic demand, and therefore may challenge urban traffic management systems significant enough to require special traffic control strategies.

Fig. 2(a) shows participation rates for diverse weekend activities. Activities that comprise more than 3% of the total are presented in the bigger pie chart on the left; others are shown in the smaller bar chart to the right. Traveling and eating, routines of daily life, absorb about 60% of total activities; entertainment, shopping, and sociality (20.4, 8.4, and 4.3%, respectively) dominate about 33% of the rest of 40% of weekend activities. The "typical weekday" activities: work (2.3%), school (1.2%), exercise and religious/civic activities (less than 2%), and out-of-town travel (0.04%) are relatively insignificant.

Fig. 2(b) shows the demographic composition of different weekend activities based on the activity type. Clearly AWNC and AWNNC participate most actively in such activities. Except for school activity, AWNCs and AWNNCs account for 60–70% of participants. The shopping activity of these two subgroups takes about 78% of the total; sociality, 64%; and out of town, 60%.

Table 2. Goodness-of-Fit Tests for Individual Activity Type

	Travel related		Work		School		Shopping		Sociality		Eating		Entertainment/ leisure		Exercise		Religious, etc.		Out-of-town	
Distribution	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR
Weibull	127.45	0.81	12.25	0.98	1.53	0.98	29.16	0.98	**2.51**[a]	**0.99**	19.41	0.98	9.15	0.99	3.06	0.98	4.45	0.96	2.53	0.84
Lognormal	**49.21**	**0.91**	41.51	0.92	4.39	0.93	15.71	0.99	19.88	0.96	9.22	0.99	37.28	0.94	4.45	0.95	10.45	0.91	2.93	0.77
Exponential	133.57	N/A	83.53	N/A	18.69	N/A	18.14	N/A	7.18	N/A	96.98	N/A	80.46	N/A	49.57	N/A	93.35	N/A	3.66	N/A
Loglogistic	58.99	0.90	37.57	0.92	3.96	0.93	23.68	0.99	20.67	0.96	10.28	0.99	33.50	0.94	3.10	0.97	7.60	0.93	2.92	0.78
Three-parameter Weibull	98.24	0.84	3.90	0.98	0.69	0.99	13.05	0.99	2.62	0.99	17.67	0.98	4.49	0.99	3.29	0.99	3.47	0.99	**2.36**[b]	**0.91**
Three-parameter lognormal	50.63	0.92	**3.10**[a]	**0.99**	**0.67**[b]	**0.99**	**10.16**	**1.00**	6.02	0.99	**8.12**	**0.99**	**2.75**[a]	**1.00**	1.49	0.99	**1.54**[b]	**1.00**	2.33	0.89
Two-parameter exponential	100.22	N/A	82.13	N/A	16.59	N/A	39.78	N/A	7.13	N/A	88.46	N/A	77.37	N/A	46.61	N/A	91.26	N/A	3.96	N/A
Three-parameter loglogistic	61.07	0.90	5.33	0.98	0.86	0.99	23.02	0.99	12.83	0.98	10.46	0.99	4.75	0.99	**0.76**[b]	**1.00**	2.08	0.99	2.30	0.90
Smallest extreme value	191.63	0.38	69.94	0.91	6.01	0.93	584.98	0.73	125.1	0.81	139.88	0.77	95.32	0.86	30.16	0.84	37.47	0.87	2.38	0.91
Normal	102.77	0.49	9.04	0.98	0.89	0.99	184.82	0.86	30.75	0.92	46.01	0.88	17.42	0.95	7.87	0.93	8.09	0.95	2.33	0.89
Logistic	99.84	0.52	10.52	0.97	0.87	0.98	177.40	0.87	34.19	0.92	43.68	0.89	16.08	0.96	6.73	0.94	8.31	0.96	2.29	0.90
Best-fit model	Lognormal		Three-parameter lognormal		Three-parameter lognormal		Three-parameter lognormal		Weibull		Three-parameter lognormal		Three-parameter lognormal		Three-parameter loglogistic		Three-parameter lognormal		Three-parameter Weibull	

Note: N/A=not applicable.

[a]Insignificant at 99% confidence level [critical AD value at the 99% confidence level is 3.857 for $n \geq 8$, Lewis (1961)].

[b]Insignificant at 95% confidence level [critical AD value at the 95% confidence level is 2.492 for $n \geq 8$, Lewis (1961)].

Based on these observations, greater attention should be given to the travel behavior of adults (19–64 years old).

Study Models and Results

As mentioned in the previous section, duration modeling is of three kinds: full parametric, nonparametric, and semiparametric. One of the first things to determine is which model kind should be used. Because the nonparametric approach does not incorporate parameters and thus cannot be used for policy analysis, it was not considered for modeling; it was, however, used to help judge whether subgroups share a common survival function (SF). The choice between the full and semiparametric approach is based on whether a proportional-hazard assumption is appropriate, and whether or not a good parametric fit can be achieved.

Based on our primary analyses (see previous section), investigation focused on activity and personal type. First, 11 parametric HFs were explored for each activity (e.g., shopping or working) and the best-fit parametric models identified. The following distributions were tested: Weibull, lognormal, exponential, loglogistic, three-parameter Weibull, three-parameter lognormal, two-parameter exponential, three-parameter loglogistic, smallest extreme value, normal, and logistic. Table 2 shows the best-fit models for 10 activity types assessed. Distributions were tested and the fitness was evaluated by means of adjusted Anderson-Darling test statistics (AD values in the table) and correlation coefficients (COR in the table). The criterion was to select the distribution with the lowest AD value or the highest COR value. As shown in Table 2, the resulting best-fit models for the 10 weekend activities are: lognormal or three-parameter lognormal for travel-related activity, work, school, shopping, eating, entertainment/leisure, religious, and civic; Weibull or three-parameter Weibull for sociality and out-of-town activities; and three-parameter loglogistic for exercise activity. The AD and COR values for the best-fit models are identified with the bold font in the table. It can be found that, in general, the best-fit models identified all have a high COR value and a low AD value. AD values below the critical values at the 95% (2.492) and the 99% (3.857) confidence level are all marked with "b" and "a," respectively. Except for travel-related and out-of-town activity, COR values for the best-fit models are mostly 0.99 or 1.00 and thus indicate a good fit. However, relatively large AD values (>8.0) for travel related, shopping, and eating activities indicate a contradictory conclusion: lack of fit and a need for a finer model. Close examination to the data of these activities revealed that most study subjects reported their durations in integer minutes (e.g., 1, 5, or 10 min) rather than "real spell" (e.g., 1.2 or 5.5 min). The reported durations are mostly clustered to these imputed values (e.g., 1, 2, or 5 min). AD statistics here, which show the squared distance between the fitted line (based on a chosen distribution) and the nonparametric step function (based on the plot points), weigh more heavily in the tail areas of the distribution (D'Agostino and Stephens 1986). Evidently the statistics are very sensitive to imputed values at the left tail of the distribution and result in large AD values.

The small number of observations resulted in a deteriorated correlation coefficient for out-of-town activity, for which there are only 15 observations. Because of nonimputed data, however, the AD statistic shows a very good fit. The AD value for the best-fit model (three-parameter Weibull) is only 2.36, which is less than the significant point of 2.492 at the 95% confidence level (Lewis 1961). The high percentage of imputed observations for travel-related activity resulted in a low correlation coefficient (0.91) and a large AD test statistic (49.2).

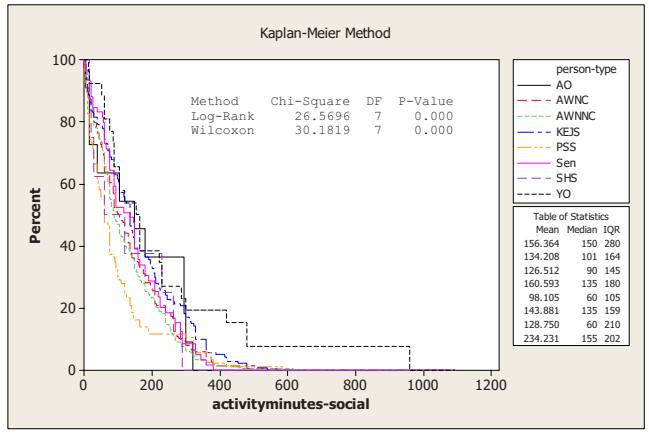

Fig. 3. Nonparametric survival plots for demographic subgroups' sociality activities

To investigate whether demographic subgroups (e.g., AWNC and SEN) for each activity could be combined and modeled with one HF/SF, the Kaplan-Meier method was used to check underlying distributions of the data. Fig. 3 provides survival plots for demographic subgroups of the sociality activities and calculated test statistics. The logrank and Wilcoxon statistics shown in the figure are significant at 99.9% confidence level with the p values of 0.000. The results support that these subgroups are significantly different and that individual HFs should be specified. Crossovers among subgroup SFs indicated that semiparametric proportional-hazard approaches are not appropriate unless these curves are approximately parallel with, rather than intersecting, each other. Again, 11 candidate parametric models were tested based on the subgroups and best-fit models specified. Table 3 shows the AD and COR values of these models. The best-fit models, shown in bold font again in the table, are those with the largest correlation coefficients and/or the lowest AD values. It can be found from the table that the best-fit models identified all have a higher COR (>0.96) and a low AD value (<1.7). These statistics indicate these models fit the data very well. Another finding from the table is that the best-fit models for subgroups are consistent with those of the aggregated group. For six out of eight subgroups, the Weibull/three-parameter Weibull was identified as the best-fit model. Other model types were chosen in only two cases: the three-parameter lognormal model for the KEJS group and the three-parameter loglogistic model for the YO group. But even in these cases, the differences between the AD and COR values of the Weibull/three-parameter Weibull and those of the identified best-fit models (three-parameter lognormal or loglogistic) are very small (less than 0.3). Therefore, the model based on the Weibull distribution can be used for every subgroup to reduce complexity in the modeling process. With continuous modeling efforts, the high correlation coefficients (>0.96) and low AD values (<1.7) show the improved goodness-of-fit. The study results emphasize that the parametric rather than the semiparametric approach should be used when underlying data distributions can be clearly identified (Hensher and Mannering 1994; Mohammadian and Doherty 2004). Analyses of the other activities also revealed that in most cases, subgroups share a common best-fit model form (albeit with different parameters).

Fig. 4 shows the estimated probability density functions (PDFs), probability plots, SFs, and HFs for sociality subgroups

Table 3. Goodness-of-Fit Tests for the Demographic Subgroups of Sociality Activities

	Assumed distributions																					
	Normal		Exponential		Two-parameter exponential		Weibull		Three-parameter Weibull		Lognormal		Three-parameter lognormal		Smallest extreme value		Logistic		Loglogistic		Three-parameter loglogistic	
Personal type	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD
AO	0.94	1.60	N/A	1.79	N/A	2.05	**0.96**	**1.53**[a]	0.965	1.53	0.931	1.72	0.945	1.58	0.919	1.97	0.932	1.67	0.926	1.79	0.935	1.65
AWNC	0.95	5.47	N/A	3.05	N/A	3.478	0.98	1.41	**0.989**	**0.87**[a]	0.94	6.40	0.988	1.38	0.862	24.01	0.948	6.39	0.94	6.04	0.975	2.75
AWNNC	0.91	14.41	N/A	5.25	N/A	4.676	0.99	1.35	**0.995**	**1.31**[a]	0.968	7.35	0.989	3.05	0.81	56.27	0.915	16.24	0.966	8.00	0.977	5.97
KEJS	0.96	3.22	N/A	3.84	N/A	3.846	0.98	1.58	0.985	1.46	0.938	5.48	**0.985**	**1.20**[a]	0.882	15.70	0.956	3.82	0.937	5.53	0.972	2.01
PSS	0.86	7.60	N/A	2.04	N/A	2.364	0.99	0.82	**0.994**	**0.63**[a]	0.984	1.03	0.991	0.73	0.747	24.49	0.869	6.84	0.981	1.20	0.984	1.06
Sen	0.94	1.86	N/A	1.73	N/A	1.756	0.96	1.77	**0.983**	**1.67**[a]	0.962	1.73	0.963	1.72	0.897	2.27	0.93	1.90	0.955	1.77	0.957	1.76
SHS	0.97	1.40	N/A	3.27	N/A	3.166	0.98	0.71	**0.993**	**0.52**[a]	0.928	1.87	0.988	0.70	0.906	5.83	0.963	1.65	0.932	1.84	0.978	1.02
YO	0.85	2.94	N/A	0.94	N/A	1.372	0.97	1.11	0.979	0.99	0.982	0.76	0.989	0.71	0.755	9.09	0.86	2.51	0.984	0.69	**0.989**	**0.66**[a]
Average	0.93	4.81	N/A	2.74	N/A	2.84	0.98	1.29	**0.99**	**1.12**[a]	0.95	3.29	0.98	1.38	0.85	17.45	0.92	5.13	0.95	3.36	0.97	2.11

Note: N/A = not applicable.

[a]Insignificant at 95% confidence level [critical AD value at the 95% confidence level is 2.492 for $n \geq 8$, Lewis (1961)].

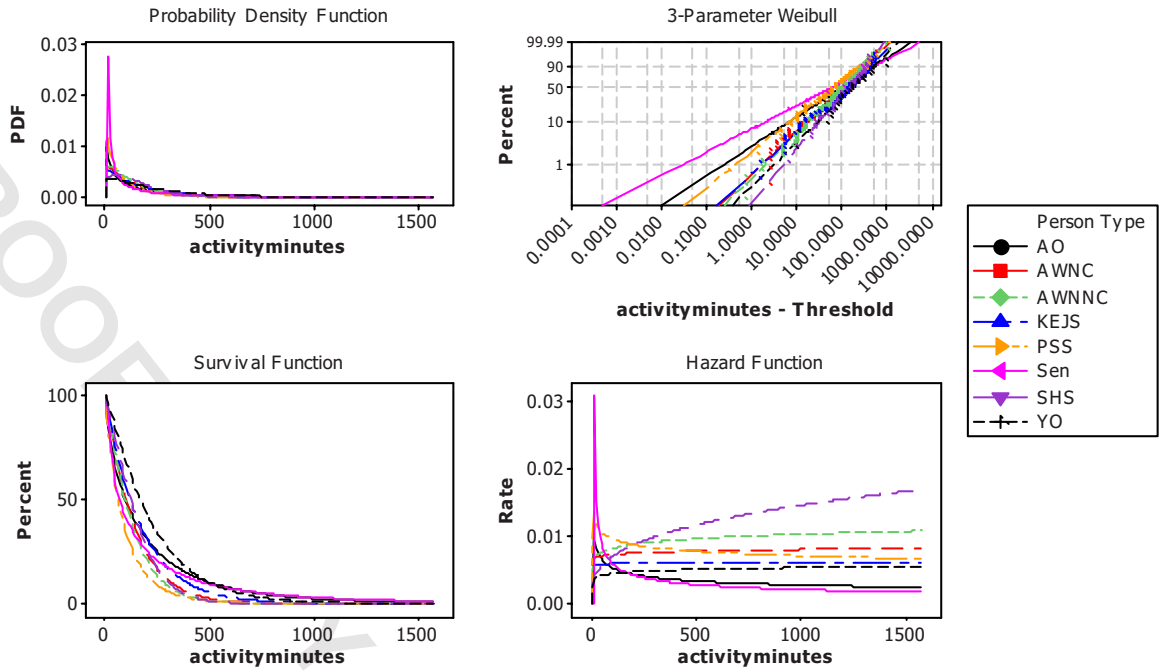

Fig. 4. Estimated PDFs, SFs, and HFs for the subgroups of sociality activities

based on three-parameter Weibull distributions. According to the probability plots, the data fit the lines very well. The PDFs, SFs, and HFs for subgroups are different from each other, especially with respect to HF. For example, the hazard rate for the AWNNC and SHS subgroups increase with duration, whereas the others have a decreased (for the SEN and PSS groups) or nearly constant (for the KEJS group) hazard rate. Close examination to the HFs indicates that hazard rates for most groups change significantly during the first 50 min, but are fairly stable thereafter.

With estimated parameters, hazard rates can be easily calculated at any time. Based on the three-parameter Weibull distribution, the estimated HF for the sociality activity of the AWNC subgroup is

$$h(t) = \lambda\beta[\lambda(t-\alpha)]^{\beta-1} = \frac{1.043}{142.09}\left(\frac{t+1.58}{142.09}\right)^{0.043}$$

where $t>0$. The hazard rate at time $t=10$ can then be calculated as 0.0066. This means that when sociality duration lasts to 10 min, the probability of abandoning such activity is 0.66% for this particular demographic group. That is, out of 1,000 people from this group, about seven will stop their sociality activity when its duration approaches to 10 min. The shape parameter $\beta=1.043$ is greater than 1 and thus indicates that the hazard is increasing as the duration increases.

Lifetime regression models were explored with a view to predict duration based on activity and socioeconomic attributes. Such models can be viewed as a starting point for developing policy-responsive duration models. Another log-LRT was also carried out for the sociality data, to test whether a full model for considering all demographic subgroups is transferable. Basically, it was found that the model is not transferable because the LRT statistic (73.5) is greater than the critical chi-square value (55.8) at the 95% confidence level with a DOF of 40. However, it can be found that there is not much difference between the LRT statistic and the calculated chi-square value, and the failure of transferability test can be attributed to the rounding of reported durations. Therefore, for the purpose of simplicity, two types of models including all demographic groups are developed within this study: one considers demographic group as a fixed effect and the other one considers it as unobserved heterogeneity.

An initial model was developed (results are not shown here) by including every independent variable (including person type, start/end time, annual income level, household size, driving capability, and age group) and that resulted in redundant information. It was found that most variables within the model were not significant at the 95% confidence level. The model resulted in a loglikelihood of −5,846.18. A kind of "stepwise regression," that involves deleting or adding independent variables one by one, was used to refine the model. The rationale behind this approach is the principle of "parsimony." During the process, an interim full model with all independent variables and the square term of the "start time" was also developed. The resulting loglikelihood was −5,832.96 and the square term of the start time was found significant. Therefore, the final models identified include the square term of the start time, start time, and the driving capability (indicated by holding a driver license or not). Fig. 5, based on the sociality data, shows the regression table for the two final models, with Fig. 5(a) for the first type of model (fixed group effect) and Fig. 5(b) for the second type (unobserved heterogeneity) mentioned earlier. Although the number of independent variables was reduced from 13 of the initial model to 3, the loglikelihood was improved from −5,846 (of the initial full model) to −5,842 [of the reduced model shown in Fig. 5(a)]. The covariates changed to become significant, as supported by the large z values and small p values. Moreover, a LRT was also carried out between the full model with the square term of the start time and the reduced model shown in Fig. 5(a). The LRT statistic was calculated as $\lambda = -2\times(-5,842.2+5,832.96)=18.48$, whereas the critical value of the chi-square distribution at the 95% confidence level with a DOF of 10, $\chi^2_{0.95}(10)$ is 18.31 (note that the reduced model has 10 fewer variables than the full model). Since the p value for the LRT statistic is 0.0474 that is very close to 5%, we can conclude that the reduced model provides almost an equal fit to the data,

```
Response Variable: activityminutes-sociality
Uncensored value         995
Estimation Method: Maximum Likelihood
Weibull distribution
```

(a) Regression Table for the Final Model without considering group heterogeneity

Predictor	Coef	Standard Error	Z	P	95.0% Normal CI Lower	Upper
Intercept	5.78655	0.168197	34.40	0.000	5.45689	6.11621
Start	-3.31782	0.600071	-5.53	0.000	-4.49394	-2.14170
Start*Start	3.19016	0.531231	6.01	0.000	2.14896	4.23135
Idrv_lic						
Y	-0.238518	0.0701853	-3.40	0.001	-0.376079	-0.100958
Shape	1.09893	0.0275447			1.04625	1.15427
Scale	177.6433					

Loglik(model)= -5842.2 Loglik(intercept only)= -5869.2

Final model McFadden's R^2= 0.004602289

Chi-sq= 54.02 on 3 degrees of freedom, p= 1.1e-11

(b) Regression Table for the Final Model considering group heterogeneity

Predictor	Coef	Standard Error	Z	P	95.0% Normal CI Lower	Upper
(Intercept)	5.7989	0.1788	32.43	1.00E-230	5.448452	6.149348
startnum	-3.3728	0.6023	-5.6	2.14E-08	-4.553308	-2.192292
startnum2	3.2297	0.5332	6.06	1.39E-09	2.184628	4.274772
IdrvlicY	-0.2312	0.0988	-2.34	1.93E-02	-0.424848	-0.037552
Shape	1.1034	0.0277			1.049108	1.157692
Scale	173.6852					

Fraity(PersonType) Chi-sq=5.36, df=2.77, p=1.3e-01

Loglik(model)= -5837.4 Loglik(intercept only)= -5869.2

Final model McFadden's R^2= 0.005419134

Chisq= 63.61 on 5.3 degrees of freedom, p= 3.4e-12

Fig. 5. Final lifetime regression models

but in a much parsimonious form. Thus the reduced model is preferred and kept.

Fig. 5(a) shows the two most important factors influencing the duration of people's sociality activity. The "driving capability" is to measure whether a person has a driver's license and thus indicates his/her mobility level. The start time used in the model is actually a number ranging from 0 to 1, with "0" representing for the midnight and 0.5 for the noon. It is used as a proxy representing how far a start time is from or to the midnight (or noon) and thus indicates a person's time constraint in executing an activity. Study results clearly confirm the importance of people's mode and time constraints in influencing their activity durations.

A lifetime regression model for considering unobserved heterogeneity at the demographic group level was also developed, as shown in Fig. 5(b). It can be found from the figure that by considering group heterogeneity, the loglikelihood was improved from −5,842 to −5,837 and the McFadden's R^2 was increased from 0.0046 to 0.0054. However, the p value for the frailty term of the regression model is not significant at the 95% confidence level. It thus indicates that the group heterogeneity has been explained by all the predictor variables and there is no unobserved heterogeneity among the groups at the social activity level. The small McFadden's R^2 values indicate that the data have a large amount of random noise. This may be attributed to the inherent rounding nature of reported duration data. However, the selected predictor variables are significant and are able to explain the observed variation in the data.

The shape parameters for the final regression models are all greater than 1.0 within their 95% confidence intervals, as shown in Fig. 5. This indicates that an increasing HF should be used for modeling sociality durations.

Concluding Remarks

Household activities and corresponding travel patterns are important themes of activity-based transportation planning. Previous research has focused on weekday activities, whereas relatively little attention has been paid to weekends (Allison et al. 2005; Lockwood et al. 2005). This study, which examined weekend household activities in Calgary, Canada, compared urban weekend versus weekday activity patterns; identified important influencing factors for activity durations; and specified best-fit hazard/

lifetime regression models based on activity type and demographics.

Pattern analyses revealed weekend activities to be substantially different from those of weekdays (Fig. 1). Most weekend household activities are executed in the afternoon. Compared to weekday activities, there are less peaking patterns for the weekend activities. For example, there is only one peak for most weekend activities such as shopping at noon, or shortly after it, whereas weekday activities usually peak twice or three times and show wider distribution. Such distinct weekend activity and related travel patterns suggest they deserve special attention, as different traffic operation and control strategies may be required to accommodate them.

The analyses carried out during this study confirm that non-work related activities predominate weekends: the total participation rate of shopping, sociality, and entertainment is over 30%, whereas work-related activity (work and school) is only about 4.5% of the total [see Fig. 2(a)]. Analyses of the demographic composition of different activities revealed that whether or not they use cars, adults 19–64 years old [AO, AWNC, and AWNNC subgroups, Fig. 2(b)] are the most active of the groups examined, comprising more than 80% of the total activities (except for "school"). Future research, therefore, should focus on the activities and related travel behaviors of them.

The applicability of parametric, nonparametric, and semiparametric duration models was closely examined. Nonparametric models, which do not allow for variables, were deemed inappropriate for policy analysis; they are, however, useful for checking underlying distributions and helpful for specifying appropriate parametric models. They were also used in this study to test whether subgroups share a common SF, and whether, therefore, proportional-hazard forms might be appropriate (Fig. 3). The semiparametric approach was not considered because (1) Kaplan-Meier plots showed that the SFs of certain subgroups overlap and hence violate the assumption of proportional hazards and (2) competing parametric models showed a high goodness-of-fit. The most frequently selected models were lognormal, followed by Weibull and loglogistic models. This research also revealed that models selected at the aggregate level (e.g., by activity type) are highly consistent with those selected at the disaggregate level (e.g., by personal type or demographic group) (see Table 3).

Kitamura (1996) mentioned that only Weibull distributions are considered for exclusively modeling durations of 18 daily activities, such as sleep, personal care, child care, meal, domestic chore, work, and work-related school and study, in a framework called prism-constrained activity-travel simulator. The results from this study, however, clearly show that may not be appropriate and a different distribution may need to be used for different activity (Table 2).

Lifetime regression models were explored for each household weekend activity. Models that included many independent variables resulted in redundant information and most of them were insignificant. For improved accuracy, a stepwise regression technique, which conforms to the principle of parsimony and potentially alleviates the future data collection burden, was used to refine the lifetime regression model. Including only three independent variables in the final models, start time, the square tem of the start time, and driving capability [see Figs. 5(a and b)], significantly improved likelihood ratios and goodness-of-fit. It was also found that the model fit will be increased by considering the group heterogeneity; however, it was not significant and thus indicates probably it is not necessary to consider such effect at the activity level.

The "rounding" or "imputed nature" of reported durations in the data resulted in deteriorated goodness-of-fit and reduced prediction power of the models developed. They would likely be more accurate if "real" durations had been reported. "Imputing" rounded observations back to a normally distributed population might solve the problem. Future research will explore this issue.

Acknowledgments

The writers are grateful to the Natural Science and Engineering Research Council, Canada (NSERC) its financial support, and to the city of Calgary for use of the data.

References

Allison, M. L., Srinivasan, S., and Bhat, C. R. (2005). "An exploratory analysis of weekend activity patterns in the San Francisco Bay Area." *Proc., 84th Transportation Research Board Annual Meeting*, Transportation Research Board, Washington, D.C.

Bhat, C. R. (2000). "Duration modeling." *Handbook of transport modelling*, D. A. Hensher and K. J. Button, eds., Elsevier Science, Oxford, UK, 91–111.

Bhat, C. R., and Lockwood, A. (2004). "On distinguishing between physically active and physically passive episodes and between travel and activity episodes: An analysis of weekend recreational participation in the San Francisco Bay Area." *Transp. Res., Part A: Policy Pract.*, 38(8), 573–592.

Bhat, C. R., and Srinivasan, S. (2005). "A multidimensional mixed ordered-response model for analyzing weekend activity participation." *Transp. Res., Part B: Methodol.*, 39(3), 255–278.

Chu, Y. L. (2005). "Modeling workers' daily non-work activity participation and duration." *Proc., 84th Transportation Research Board Annual Meeting*, Transportation Research Board, Washington, D.C.

Cox, D. R. (1972). "Regression models and life tables." *J. R. Stat. Soc. Ser. B (Methodol.)*, 34, 187–220.

D'Agostino, R. B., and Stephens, M. A. (1986). *Goodness-of-fit techniques*, Marcel Dekker, New York.

Doherty, S. T., Lee-Gosselin, M., Burns, K., and Andrey, J. (2002). "Household activity rescheduling in response to automobile reduction scenarios." *Transp. Res. Rec.*, 1807, 174–181.

Ettema, D., Borgers, A., and Timmermans, H. (1995). "Simulation model of activity scheduling behaviour." *Transp. Res. Rec.*, 1413, 1–11.

Federal Highway Administration (FHWA). (2004a). "Traffic congestion and reliability: Linking solutions to problems." *FHWA, U.S. DOT*, ⟨http://www.ops.fhwa.dot.gov/congestion_report_04/congestion_report.pdf⟩ (Dec. 6, 2006).

Federal Highway Administration (FHWA). (2004b). "2001 national household transportation survey." *FHWA, U.S. DOT*, ⟨http://nhts.ornl.gov/2001/pub/STT.pdf⟩ (Nov. 18, 2006).

Gärling, T., Kwan, M. P., and Golledge, R. G. (1994). "Computational-process modelling of household activity scheduling." *Transp. Res., Part B: Methodol.*, 28(5), 355–364.

Gavaghan, D. J., Moore, A. R., and McQay, H. J. (2000). "An evaluation of homogeneity tests in meta-analysis in pain using simulations of patient data." *Pain*, 85, 415–424.

Gilbert, C. (1992). "A duration model of automobile ownership." *Transp. Res.*, 26(2), 97–114.

Hamed, M. M., and Easa, S. M. (1998). "Integrated modeling of urban shopping activities." *J. Urban Plann. Dev.*, 124(3), 115–131.

Hamed, M. M., and Mannering, F. L. (1993). "Modeling travelers' post-work activity involvement: Toward a new methodology." *Transp. Sci.*, 27(4), 381–394.

Hensher, D. A., and Mannering, F. L. (1994). "Hazard-based duration models and their application to transport analysis." *Transport Rev.*,

14(1), 63–82.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). "Measuring inconsistency in meta-analyses." *BMJ*, 327, 557–560.

Hunt, J. D., and Atkins, D. M. (2004). "Characteristics of weekend travel in Calgary." *Proc., Transportation Revolutions: Proc. of the 39th Annual Conf. of the Canadian Transportation Research Forum*, Canadian Transportation Research Forum (CTRF), Calgary, Canada.

Hunt, J. D., Stefan, K., and McMillan, P. (2005). "Microsimulating weekend travel by households in Calgary." *Proc., 2005 Transportation Association of Canada Conf.*, Transportation Association of Canada (TAC), Calgary, Canada.

Kitamura, R. (1996). "Applications of models of activity behavior for activity based demand forecasting." *Proc., Activity-Based Travel Forecasting Conf.*, USDOT, New Orleans.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data*, Wiley, Hoboken, N.J.

Lewis, P. A. W. (1961). "Distribution of the Anderson-Darling statistic." *Ann. Math. Stat.*, 32(4), 1118–1124.

Lockwood, A. M., Srinivasan, S., and Bhat, C. R. (2005). "An exploratory analysis of weekend activity patterns in the San Francisco Bay Area." *Proc., 84th Transportation Research Board Annual Meeting*, Transportation Research Board, Washington, D.C.

Mannering, F. L., and Hamed, M. M. (1990). "Occurrence, frequency, and duration of commuters' work-to-home departure delay." *Transp. Res., Part B: Methodol.*, 24(2), 99–109.

Miller, E. J., and Roorda, M. J. (2003). "Prototype model of household activity/travel scheduling." *Transp. Res. Rec.*, 1831, 114–121.

Mohammadian, A., and Doherty, S. T. (2004). "A hazard model for the duration of time between planning and execution of an activity." *Proc., Conf. on Progress in Activity-Based Analysis*, International Conference on Recent Advances in Retailing and Services Science (EIRASS), Maastricht, The Netherlands.

Nam, D., and Mannering, F. (2000). "An exploratory hazard-based analysis of highway incident duration." *Transp. Res., Part A: Policy Pract.*, 34, 85–102.

Parsons Brinckerhoff (PB) Quade and Douglas, Inc. (2000). "Comparative analysis, weekday and weekend travel with NPTS integration for the RT-HIS: Regional travel-household interview survey." *Rep. Prepared for the New York Metropolitan Transportation Council and the North Jersey Transportation Planning Authority, Inc.*

Paselk, T. A., and Mannering, F. L. (1994). "Use of duration models for predicting vehicular delay at a US/Canadian border crossing." *Transportation*, 21, 249–270.

Sall, E. A., Bhat, C. R., and Reckinger, J. (2005). "An analysis of weekend work activity patterns in the San Francisco Bay Area." *Proc., 84th Transportation Research Board Annual Meeting*, Transportation Research Board, Washington, D.C.

Stathopoulos, A., and Karlaftis, M. G. (2002). "Modeling duration of urban traffic congestion." *J. Transp. Eng.*, 128(6), 587–590.

Wang, W. Q., Chen, H. B., and Bell, M. (2002). "A study of the characteristics of traffic incident duration on motorways." *Proc., 3rd Int. Conf. on Traffic and Transportation Studies*, ASCE, Reston, Va., 1101–1108.

Washington, S., Karlaftis, M., and Mannering, F. (2003). *Statistical and econometric methods for transportation data analysis*, CRC, Boca Raton, Fla.

PROOF COPY [TE/2007/023759] 002003QTE