Duration Modeling of Calgary Household Weekday and Weekend Activities: How Different Are They?

By

Ming Zhong*

Assistant Professor
Dept. of Civil Engineering, University of New Brunswick
GD-128, Head Hall, 17 Dineen Drive
P.O. Box 4400, Fredericton, N.B., Canada E3B 5A3
Telephone: (506) 452-6324 (O) Fax: (506) 453-3568
Email: ming@unb.ca

John Douglas Hunt

Professor
Institute for Advanced Policy Research, University of Calgary
2500 University Drive N.W., Calgary, Alberta, Canada T2N 1N4
Phone: (403) 220-8783 Fax: (403) 282-7026
Email: jdhunt@ucalgary.ca

And

Xuewen Lu

Assistant Professor
Department of Mathematics and Statistics
University of Calgary, Calgary, AB. T2N 1N4 Canada

Phone: (403) 220-6620 Fax: (403) 282-5150
Email: lux@math.ucalgary.ca

Paper prepared for presentation

at the Best Practices in Urban Transportation Planning - An Economic Enabler Session

of the 2007 Annual Conference of the
Transportation Association of Canada
Saskatoon, Saskatchewan

---

*Corresponding author

Duration Modeling of Calgary Household Weekday and Weekend Activities: How Different Are They?

By Ming Zhong, John Douglas Hunt, and Xuewen Lu

Abstract

A desirable activity-based travel demand modeling framework should be able to address both weekday and weekend activities. However, a literature review shows previous research efforts have mostly focused on investigating weekday but not weekend activities. Little or no research exists to quantify the differences between weekend and weekday activities. The best knowledge to date is limited to weekday and weekend activities starting at different time of the day and with different participation rates. This study aims to fill the gap by studying the differences between weekday and weekend activities in Calgary, Canada, in terms of their participation rates, starting time, duration and inferred location choices. First, statistics of these attributes were computed for 10 types of weekday and weekend activities and they were found different. Secondly, Log-rank and Wilcoxon tests further proved a common type of weekday and weekend activity tend to follow different survival functions. Third, best-fit duration models were explored for each type of weekday and weekend activity and compared with each other. It was found that Lognormal and Weibull were chosen as the best-fit models for nearly all weekday and weekend activities. The best-fit duration models for same types of weekday and weekend activities (e.g., shopping) were different in either underlying distribution or estimated parameters. This study clearly shows the weekend activities are different from their weekday counterparts and suggests that they should be treated separately in activity-based modeling frameworks (237 words).

Introduction

Activity-based travel demand modeling is relatively new and has not been widely used in practice [1-2]. It has many advantages over traditional trip-based approach, such as richness, attractive theoretical elegance, and intuitive implementation with micro-simulation. It, however, requires significantly more financial and personal resources for collecting data and carrying out detailed analysis. Advances in information technology, such as object-oriented database design, high-performance computing, and geographic information systems (GIS), have made onerous data collection, management, and processing easy, and the model development cost has been continuously reduced. All of these encourage more such applications in transportation areas. Within activity-based modeling framework, duration modeling has stood out as one of important themes as it is an inherent nature of any activity. Once activity sequence, duration, and location choices are determined, travel can be captured as the "induced" demand raised by other activities.

A literature review indicates that many studies have been carried out for investigating weekday activities and related travels [3-5], but less effort has been placed on weekends [6-9]. The logic may be that high travel demand at morning and evening peak hours on weekdays result in frequent traffic congestion and thus warrants special attention. However, as travel demand increases and infrastructure construction is constrained, traffic congestion takes place in recreational areas, major shopping centers, sports arenas, and bridges of many big cities over weekends [10]. Moreover, real-time traffic operation and management offered by Intelligent Transportation Systems (ITS) requires travel demand information on weekends. As such, travel demand modeling on weekdays and weekends is required at both the planning and operation levels. It is desired to integrate them into one modeling framework. This challenges traditional approaches of analyzing weekday and weekend activities separately and warrants a study to investigate them together.

Within the existing literature of modeling weekday and weekend activities, none has been found to study the differences of the two. The knowledge is generally limited to that weekday and weekend activities have different participation rates and starting time [11]. Therefore, this study is aimed to empirically quantify the difference of weekday and weekend activities, in terms of their activity type, starting time, duration, and inferred location choices. The best-fit duration models for individual weekday and weekend activities are also explored and compared.

In this study, a general picture of 10 weekday and weekend household activities is presented first. Comparisons are made in terms of their participation rates, starting time, and durations. Secondly, the Kaplan-Meier method is used to test whether a same type of weekend/weekday activity would follow the same survival function. The analyses are then continued by choosing the best-fit hazard functions for each type of weekday/weekend activity. For each weekday and corresponding weekend activity, the best-fit models are specified and compared with each other.

The rest of the paper is organized as the following. First, a literature review for duration models in general and their applications for weekend/weekday activity modeling in particular is presented; then the data used in this study are briefly introduced; study models and results are discussed; and finally, major findings and conclusions are given.

Literature Review

Review of duration models

The statistical analysis of what are called as lifetime, survival time, or failure time data is known as duration modeling or hazard modeling, and has long been an important topic in many areas, such as biomedical, engineering, and social sciences [12]. For example, Bartholomew [13] used duration models to study the lifetime distribution of equipment. Prentice [14] compared the effects of two chemotherapy treatments in prolonging survival time of 40 advanced lung cancer patients. Nevertheless, duration models have been increasingly used recently in transportation area [1-2].

The key element of duration modeling is hazard functions, which indicate the way that the risk or probability of failure varies with age or time [12]. A hazard function $h(t)$ that expresses the probability of the occurrence of an event during a certain time interval, say t to $t + \Delta t$, given that the event has not occurred or ended before the beginning of the interval. The conditional probability of duration starting or ending plays an important role as the probability indicates that an event starts or terminates depends on the length of time or the duration has lasted [12].

Let T be a nonnegative random variable representing the lifetime of individuals in some population. Here only a continuous variable T is assumed (discrete T can be accommodated by considering the discretization as a result of segmentation of continuous time into several discrete intervals), as this is the case for most applications. Let the probability density function (PDF) of T is $f(t)$ and the cumulative distribution function (CDF) is $F(t)$. Then we have:

$$F(t) = P(T < t) \tag{1}$$

Where t is some specific time period. In the case of household activities, the CDF is defined to indicate the probability of an activity would last less than a specified time period, t. Then the PDF, $f(t)$ can be obtained:

$$f(t) = \frac{dF(t)}{dt} \tag{2}$$

This provides unconditional distribution of duration T. The survival function (SF), $S(t)$, is then can be defined as:

$$S(t) = P(T \geq t) = 1 - F(t) \tag{3}$$

The above survival function represents the probability that the duration in a state T will be greater than or equal to the specific time t. The hazard function can then be expressed as a function of the PDF $f(t)$, the CDF $F(t)$ and SF $S(t)$, as shown in the following equation:

$$h(\mathrm{t}) = \lim_{\Delta t \to 0} \frac{p(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(\mathrm{t})}{1 - F(\mathrm{t})} = \frac{f(t)}{S(t)} \tag{4}$$

The hazard, $h(t)$, gives the rate at which events (e.g., ending an activity) are occurring at time t, given that the events have not occurred up to time t [12].

Parametric models

Both proportional hazard and accelerated time type of parametric models have been used in the analysis of lifetime data [1, 12]. 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 hazard function for T [12]. Among these models, a few distributions demonstrate their significance and have been applied in a wide range of situations [12]. These distributions are Exponential, Weibull, Log-logistic, Log-normal, extreme value (Gumbel) and Gamma. For a complete discussion about parametric duration models, please see Lawless [12].

Semi-parametric models

When there is no clear choice concerning hazard functions, another safer approach is to use semi-parametric hazard models [12]. For these models, there is no distributional assumption for the baseline hazard and leave it arbitrary, but assumption is made concerning the functional form specifying how the external covariates interact with the baseline hazard in the model. The semi-parametric models partially relax the assumption of parametric relationship between various factors and resulting hazard rate. However, it should be noted that the semi-parametric models stand on strong assumptions regarding how external covariates interact with the baseline hazards, and require careful checking in applications [12]. Moreover, if the hazard is generated from a known distribution and a semi-parametric model is applied, statistical efficiency will be lost since information regarding the hazard's distribution is not being used. This could result in less precise coefficient estimates as reflected by their higher standard errors [2].

Nonparametric models

The above parametric or semi-parametric hazard functions could be applied to questions with sound theoretical foundation. However, in some cases, if little or no knowledge of the functional form of the hazard is available, one might consider to use a non-parametric approach, for which there is no assumption involved concerning the underlying distribution of the baseline hazard [12]. The most popular nonparametric method is the Kaplan-Meier Estimator. For such an approach, the duration scale is split into small discrete periods and by assuming a constant hazard within each period, one can then estimate the continuous-time step function hazard shape. The Kaplan-Meier Estimator is particularly good in situations where there are a small number of groups and we want to examine whether they have similar survival distributions.

Several statistics, such as Log-Rank and the Wilcoxon, are usually used in such a test. The Log-Rank test is a form of Chi-square test, which calculates a statistic for testing the null hypothesis that the survival curves are the same for all groups. The Log-Rank test statistic is often *approximated* by the following equation:

$$\chi^2(1) = (O_A - E_A)^2 / E_A + (O_B - E_B)^2 / E_B \tag{5}$$

Where O_A, E_A, O_B, E_B are the observed & expected number of the events in group A and B respectively. The expected number of occurred events is calculated as, for group A at time i:

$$E^i{}_A = \frac{Number\ of\ events\ at\ risk\ at\ time\ i}{Total\ number\ of\ events\ at\ risk\ at\ time\ i} \times the\ number\ of\ events\ occurred\ at\ time\ i \quad (6)$$

Breslow [15] has shown the Log-Rank test is appropriate for survival functions with a "proportional hazard", that is, the relative hazard does not change with time. When the hazard rates change with time, the Wilcoxon test is more appropriate [15]. The Wilcoxon statistic is to compute the sum of how many times the lifetime observations in one group are larger than those in another one. For two groups with a size of m and n, the total number of matches is $m \times n$. If the two samples are assumed to have a same lifetime distribution, then it is expected that the times of lifetime observations in one group exceed those in the other group should be half of this number, that is, it should be equal to or about $\frac{m \times n}{2}$. A non-parametric test, which uses re-sampling techniques to calculate a p-value based on observed values from the group, is usually used to test if the null hypothesis that the two samples are from a same population is true [16].

The nonparametric shape obtained from the Kaplan-Meier Estimator is also useful for empirically testing the assumed parametric baseline shapes [12]. However, analyst cannot incorporate any explanatory variable for policy analysis with such an approach.

Duration modeling of weekend and weekday activity patterns

A review of the literature indicates that limited efforts have been contributed to model activity durations, even it is an integrated part of activity-based modeling frameworks. For example, Kitamura et al. [17] 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 PCATS. Recent developments in the area consist of proposing a general framework for modeling workers or commuters' activity and travel patterns [5, 18-20], but none of them explicitly modeled the durations of considered activities. The consequence of little research contribution in this area led to that the practitioners used observed duration distributions in their models [21].

Among the scare literature available for weekend activity duration modeling, a group of scholars at the University of Texas at Austin stand out and contributed a few studies [6-9, 22]. In particular, Lockwood et al. [22] provided a comprehensive exploratory analysis of nine categories of weekend activities, including average frequency and durations, time of day of travel, model of travel by trip purpose, trip distance by purpose, total volume of travel by trip purpose, sequencing of activity episodes, activity episode chaining, and activity purpose of the first and last out-of-home episode of the day. They basically presented a comprehensive view of weekend activities in the Bay Area by a series of statistical analyses, but no attempts were made to specify duration models for individual activities. Although Bhat and Srinivasan [8] and Sall et al. [9] outlined the methodologies of duration modelling in their weekend activity analysis framework, no statistical results were provided.

There are a number of recent studies applying duration models for non-work activities (may not be weekend activities). Ettema et al. [23] applied competing risk hazard models for modeling activity choice, timing, sequencing, and duration of 39 students at the Eindhoven University of Technology, Netherlands. Chu [24] used a Type II Tobit model to model workers' daily non-work activity durations with 1997/98 New York household survey data. Hamed and Mannering [25] estimated travel time from work to home and activity durations with ordinary least squares regression and three-stage least squares regression. The corrected R^2 of 0.11 was reported for the ordinary least squares regression and 0.188 for the three-stage least squares regression. Mannering and Hamed [26] used a Weibull-based duration model for estimating commuters' work-to-home departure delay time in Seattle. The choice of the Weibull distribution is based on their finding that that the end of a departure delay can be viewed as being induced by any one of a number of random factors, such as decrease in homeward traffic congestion, boredom with the activity undertaken, completion of activity undertaken, etc. They argued that, since the end of departure delay depends on the shortest time to the occurrence of one of these random factors, it should follow a distribution of the smallest extreme and therefore the Weibull distribution is appropriate [26]. They achieved a standard error of 0.148 for the duration parameter estimates.

Duration modeling has also been widely applied to other transportation applications. For example, Nam and Mannering [27] used hazard-based duration models to evaluate the time of incident detection/reporting, response, and clearance in Washington State. Paselk and Mannering [28] used Log-logistic duration models to predict vehicular delay at a US/Canadian border crossing. They found that the duration models are better alternatives for traditional queuing analysis as it is not easy to capture duration dependence effect with such a tool [28]. Stathopoulos and Karlaftis [29] examined four most widely used hazard functions for modeling congestion durations in Athens, and found that the Log-logistic form is most appropriate. Hensher and Mannering [2] provided a comprehensive review for the applications of hazard-based duration models in transport analysis. Lawless reviewed many applications other than transportation [12]. For a complete review of duration models and their applications, please consult this literature.

Study Data and Primary Analyses

A large scale of household activity survey was completed in the City of Calgary, Canada in late 2001 and early 2002 [30]. The purpose of the survey was to collect data for both short-term traffic operational analysis and long-term transportation planning. The data provides excellent opportunities for analyzing weekday and weekend activities and related travel behaviors, and is expected to provide insights for future policy analysis.

The data include detailed socioeconomic information of the person being surveyed and attributes of executed activities, such as personal type, employment status (fulltime or part-time), annual income level, gender, age, household size (persons in household), driving capability (licensed or not), activity type, activity durations (in minutes), and starting & ending time of each activity. The 10 activities investigated in this study include A) travel related activity (e.g., dropping off or picking up a person), D) working, E) schooling, F) shopping, I) sociality, J) eating, K) entertainment, M) exercise,

N) religious activities, and Z) out-of-town travel. The activity types and their labels were defined during the survey and they are used here directly. Approximately 13,000 weekday and weekend observations are used in this study.

Figure 1 shows the weekend (a) and weekday (b) household activity patterns. It is clear from Figure 1(a) that, except for the Religious and Civic activity, all of the other weekend activities start in the afternoon (after the middle point of 0.5 – the noon) and there is usually only one afternoon peak for the weekend activities. For example, the weekend shopping, sociality and entertainment/leisure activities show exactly such patterns. In contrast, Figure 1(b) shows that, in general, weekday activities start in the morning and there are both morning and afternoon peaks of work and school related activities. Comparison made between Figure 1(a) and 1(b) indicates that the amount of individual activities is consistent with the expected weekly periodicity. For example, there are much more work and school related activities during weekdays, but less shopping, religious, sociality, entertainment activities, whereas the weekends show the reversed patterns. The weekend activity patterns indicate that a corresponding high traffic demand will be introduced over a short period in the afternoon. This could potentially challenge weekday-based urban traffic management systems and a distinct traffic operation strategy may be required.

Figure 2 show (a) weekday and (b) weekend activity participation rates. It is clear from the comparison of the two that there are more travel and work/school-related activities during the weekdays, whereas the weekends are distinguished with more shopping and sociality/leisure-related activities. For example, the total participation rate for work and school during the weekdays is 13.8%, but it is reduced to 3.5% on the weekends. The total participation rate for typical weekend activities (including shopping, sociality, and entertainment/leisure) is only 21.5% during the weekdays, but it is increased to over 33% on the weekends. The activity participation rates shown here imply different location choices between weekdays and weekends, as people would have to travel to different destinations to fulfill their goals.

Table 1 compares the mean, median and the 75th percentile durations for the same types of weekday and weekend activities. The difference is evaluated by a percent error (PE) calculated using the following equation:

$$PE = \frac{weekend\ parameter - weekday\ parameter}{weekday\ parameter} \times 100\% \tag{7}$$

The results in Table 1 show that weekday travel, school, and out-of-town activities usually last longer than their weekend counterparts, but all the other activities tend to have shorter durations. The differences are usually within 20% of each other for corresponding weekday and weekend Travel, Work, School, Shopping, Eating, Entertainment/Leisure (Ent/Leis), Exercise activities. However, there are much larger differences for Sociality, Religious, Civil etc., and Out-of-town activities. The differences are all more than 30-40%, with some as high as 76%. Such large differences emphasize these activities are typical weekend types. Another finding from Table 1 is that there are large differences between the mean levels of durations of different activities. For example, the mean duration of travel related activities (drop-by etc.) is 18 minutes, whereas working activities on average last for over 200 minutes. The large differences between the mean and median durations of different activities emphasize that they should be modeled separately.

Study Results

The statistical analyses in the above section show that weekday and corresponding weekend activities are different in terms of their starting time, duration, participation rates and consequently location choices. This may imply that an activity-based modeling framework needs address them separately, and thus would unavoidably increase modeling complexity. Therefore, an interesting question raised with this issue is: whether weekday and weekend activities follow similar underlying distributions and can they be combined and modeled together? The section focuses on this question.

First, Kaplan-Meier tests are used to check whether corresponding weekday and weekend activity follow a same survival duration function. The method used in this study involves visually comparing the survival functions of the same type of weekend and weekday activity and calculating the following two statistics: Log-Rank and Wilcoxon. Figure 3 shows such a test for comparing weekday and weekend shopping durations. It can be seen that the two patterns meet each other quite well at the each end, but have significant differences in the middle range, especially from 30 to 200 minutes. The weekend pattern clearly shows a longer duration than its weekday counterpart. The Log-Rank and Wilcoxon statistics shown in Figure 3 also emphasize that the two distributions are significantly different at the 95% confidence level, as the p-values are all less than 0.05. Table 2 lists the Log-Rank and Wilcoxon statistics for all weekday and weekend activity pairs. The Chi-square statistic for either the Log-Rank or the Wilcoxon test is significant at 95% confidence level in 9 out of 10 cases. It is interesting to note that the weekday and weekend working durations are not significantly different indicating they may be combined in future modeling exercises.

The above analyses show that, in general, pairs of weekday and weekend activity durations distribute differently and suggest that they should be modeled separately (except for the working activity). The next task is then to identify the best-fit duration models for these activities and to see if they share a same type of best-fit model. The analyses are done by fitting weekday and weekend activity durations individually and then comparing them. Eleven parametric duration models available from MINITAB are fitted against each weekday and weekend activity, and the best-fit parametric models are identified. The following distributions are tested: Weibull, Lognormal, Exponential, Log-logistic, 3-parameter Weibull, 3-parameter Lognormal, 2-parameter Exponential, 3-parameter Log-logistic, Smallest extreme value, Normal, and Logistic. The distributions are tested and the goodness-of-fit is evaluated with adjusted Anderson-Darling test statistics (AD values in the table) and correlation coefficients (COR in the table). The criterion is to select the distribution with the lowest AD value or the highest COR value [31].

Table 3 shows the analysis results. For illustration purpose, the AD and COR values for the best-fit models are highlighted with bold fonts. To facilitate the presentation, the results for weekday and weekend activities are color-coded with red and black respectively. In general, it can be seen that a relatively high-level of goodness-of-fit is achieved, which is shown by universally high COR Values greater than 0.98. The AD values vary greatly from as low as less than 1 to over 500, depending on which distribution is selected and how many observations are available for fitting. In general, AD values greater than 2.5 indicate a lack of fit [32] and they are observed for nearly half cases. Detailed analysis indicates that this is resulted from imputed durations, which mostly cluster to certain integer spells (e.g., 1, 5, or 10 minutes). These highly imputed durations result in large AD values as Anderson-Darling tests emphasize more on the fit of the tails [31].

The resulting best-fit models selected for the 10 ***weekend*** activities are (as indicated by AD and COR values coded by black bold fonts in Table 3): 3-parameter Lognormal for Travel Related activity, Shopping, and Eating; Weibull or 3-parameter Weibull for Work, School, Sociality, Entertainment/Leisure, and Out-of-town activity; and Log-logistic or 3-parameter Log-logistic for the Exercise and Religious etc. activity, as shown in Table 3. The Weibull, especially 3-parameter Weibull, are identified as most applicable model, followed by Lognormal and Log-logistics.

The best-fit models selected for the 10 ***weekday*** activities are shown in Table 3, as indicated by AD and COR values coded with the red bold fonts: Lognormal or 3-parameter Lognormal for Travel Related activity, Work, School, Eating and Religious etc. activity; Weibull or 3-parameter Weibull for the remaining activities, which include Sociality, Entertainment/Leisure, Exercise, and Out-of-town. In contrast to the results for the weekend activities, only Lognormal and Weibull are judged as best-fit models.

The best-fit models selected above are further proved by visually comparing the fitted lines with observed cumulative distribution functions (CDFs). Figure 4(a) shows such an example for the weekend shopping activity. The 3-parameter Lognormal model fits the observed durations very well (with a COR value of 0.99, Table 3). Figure 4(b) shows the probability density function (PDF), the probability plot, survival function (SF), and hazard function (HF) based on the selected 3-parameter Lognormal model. The data shows that most people have a shopping duration less than 75 minutes, as the area under the PDF greater than this value is very small and the survival rate is very low (less than 20%). The highest hazard rate for the activity is found about 0.03 at 25 minutes. The HF shows a rapid increasing trend from the starting point to the peak and a fairly flat diminishing rate after that point. Such a HF pattern indicates that people with a shopping duration of less than 25 minutes are more sensitive to the time spent because the hazard of abandoning the activity increases very quickly as time elapses. In contrast, those who stay longer than 25 minutes tend to stay longer as their hazard rates begin to drop. The probability plot on the same figure also indicates a good fit between the theoretical function and the observed durations. However, a noticeable deviation at the lower left corner can be observed. Close examination of the data indicates it is resulting from the "imputed nature" of reported durations, as people tend to declare their durations with integer values, such as 1, 2, 5, or 10 minutes.

Attempts for testing semi-parametric models are abandoned because the competing parametric models are found to have a high degree of fit. In such cases, parametric approaches are preferred since they are more accurate and efficient, and they are better understood by practitioners.

Concluding Remarks

The credibility of transportation planning has been significantly improved through detailed study of household activities and induced travel patterns. Previous research has been focused on weekday activities, but little emphasis has been placed on their weekend counterparts [21]. This paper compares the weekday and weekend household activities in the context of the City of Calgary, Canada. The attempts made in this study include comparing the household weekday and weekend activity patterns in terms of their participation rates, starting time, duration, and "inferred" location choice; and specifying best-fit parametric models for individual weekday and weekend activities.

Pattern analyses show that weekend activities are quite different with their weekday counterparts. There is only one peak for most of weekend activities (e.g., shopping) and the peak is usually around or after the noon. On the contrary, most weekday patterns show two peaks, one in the morning and the other in the afternoon. Such a finding confirms the distinction between weekday and weekend travel patterns and suggests different traffic operation strategies may have to be applied.

A comparison between weekday and weekend activity participation rates shows that there are more travel and work/school-related activities during the weekdays, whereas, in contrast, the weekends are distinguished with more shopping and sociality/leisure-related activities. For example, the participation rate for non-work/school related activities (including Shopping, Sociality, and Entertainment/Leisure) is only 21.5% of the total during the weekdays, but it is increased to over 33% on the weekends. The different activity participation rates over weekdays and weekends imply shifts in people's location choices and thus their travel patterns. Significant increases in non-work/school activities during the weekends suggest we need pay more attention to related travels.

Descriptive statistics, such as mean, median, and the 75th percentile, are used to evaluate the differences between weekday and weekend duration patterns. It is found that there are large differences. In general, weekday School, Travel Related and Out-of-town activities are longer than those executed over the weekends. On the contrary, the other activities tend to last longer during the weekends.

The fact that durations of weekday activities and their weekend counterparts are different is further proved with Kaplan-Meier non-parametric tests. The Log-Rank and Wilcoxon test statistics are significant at the 95% confidence level in most case, which reject the hypothesis that the weekday activities and their weekend counterparts follow same duration distributions.

Eleven parametric duration models are used to fit the 10 types of weekday and weekend activities. It is found, in general, a high-level of goodness-of-fit is achieved across all types of activities. The best-fit models identified for the weekday and weekend activities are 3-parameter Lognormal, Weibull/3-parameter Weibull, or Log-logistics/3-parameter Log-logistics. The obtained COR values for the best-fit models are all equal or more than 0.98. However, the large AD values for half cases indicate a lack of fit. Detailed analysis indicates that they are resulted from the "imputed nature" of reported durations.

The duration models developed here will be used in the Calgary household modeling framework. The aim is to replace previous "static" duration models based on weighted samples from observed durations. The models developed in this study explicitly consider many characteristics of individuals and activities, but do not incorporate those variables by which some important policy analyses can be made, such as transit fare and waiting time. For example, if the city government could provide transit service with much lower fare and less waiting time, we would assume people would travel more for visiting more locations and thus they would stay for a shorter period than before at any given location. Currently, this is limited by the data used. Future research could include these variables into the analysis once such information is available. Another issue identified during this research is the "rounding" of reported durations. These "imputed" durations result in deteriorated fit. It is expected that the developed models could be more accurate if "real" durations would have been reported in the data. Such a problem may be solved by "imputing" these rounded observations back into a normally distributed population. Future research is going to explore such an issue.

Acknowledgments

The authors are grateful towards Natural Science and Engineering Research Council (NSERC), Canada for their financial support, and the City of Calgary for the data used in this study.

References

1. Bhat, C.R. (2000). Duration Modeling. *Handbook of Transport Modelling*, pp. 91-111, Edited by D. A. Hensher and K.J. Button, Elsevier Science.
2. Hensher, D.A. and F.L. Mannering. (1994). Hazard-Based Duration Models and Their Application to Transport Analysis. *Transport Reviews*, 14, pp. 63-82.Hunt, J.D. and Atkins, D.M., 2004, Characteristics of Weekend Travel in Calgary. Transportation Revolutions: Proceedings of the 39th Annual Conference of the Canadian Transportation Research Forum, Calgary AB, Canada, May 2004, pp. 253-267.
3. Doherty, S. T., M. Lee-Gosselin, K. Burns and J. Andrey. (2002). Household Activity Rescheduling in Response to Automobile Reduction Scenarios. *Transportation Research Record 1807*, pp. 174-182.
4. Gärling, T., M. P. Kwan and R. G. Golledge. (1994). Computational-process modelling of household activity scheduling. *Transportation Research B* 25(5), pp. 355-364.
5. Miller, E. J. and M. J. Roorda. (2003). A Prototype Model of Household Activity/Travel Scheduling. *Transportation Research Record 1831*, pp. 114-121.
6. Allison, M.L., S. Srinivasan, and C.R. Bhat. (2005). An Exploratory Analysis of Weekend Activity Patterns in the San Francisco Bay Area. Presented at *the 84th Transportation Research Board Annual Meeting*, Washington DC.
7. 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. *Transportation Research Part A*, Vol. 38, No. 8, pp. 573-592.
8. Bhat, C.R., and Srinivasan, S. (2005). A Multidimensional Mixed Ordered-Response Model for Analyzing Weekend Activity Participation. *Transportation Research Part B*, Vol. 39, No. 3, pp. 255-278.
9. Sall, E.A., C.R. Bhat, and J. Reckinger. (2005). An Analysis of Weekend Work Activity Patterns in the San Francisco Bay Area. Presented at *the 84th Transportation Research Board Annual Meeting*, Washington D.C.
10. Federal Highway Administration (FHWA). (2004). Traffic Congestion and Reliability: Linking Solutions to Problems. FHWA, US Department of Transportation (website: http://www.ops.fhwa.dot.gov/congestion_report_04/congestion_report.pdf, last accessed Dec. 06, 2006)
11. Hunt, J.D., McMillan, P., Stefan, K. and Atkins, D.M. (2005). Nature of Weekend Travel by Urban Households. Presented at the 2005 Annual Conference of the Transportation Association of Canada, Calgary, Canada, September 2005.
12. Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, Inc., Hoboken, New Jersey.
13. Bartholomew, D.J. (1957). A Problem in Life Testing. *Journal of American statistics Association*, 52, 350-355.
14. Prentice, R.L. (1973). Exponential Survival with Censoring and Explanatory Variables. *Biometrika*, 60, pp. 279-288.

15. Breslow, N.E. (1970). A generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship. *Biometrika*, 57: 579-594
16. Wilcoxon, F. (1945). Individual Comparison by Ranking Methods. *Biometrics*, 1, pp. 80-83.
17. Kitamura, R. (1996). Applications of Models of Activity Behavior for Activity Based Demand Forecasting. Presented at Activity-Based Travel Forecasting Conference, June 2-5, 1996
18. Arun, R.K. and Pendyala, R.M. (2001) A Structural Equations Analysis of Commuters' Activity and Travel Patterns. Transportation 28, pp. 33-54.
19. Bhat, C.R. (2001). Modeling the Commute Activity-Travel Patterns of Workers: Formulation and Empirical Analysis. Transportation Science, Vol. 35, No. 1, pp. 61-79.
20. Bhat, C.R. and Singh, S.K. (2000). A Comprehensive Daily Activity-Travel Generation Model System for Workers, Transportation Research, Part A, 34, pp. 1-22.
21. Zhong, M. and Hunt, J.D. (2005). Modeling the Durations of Calgary Household Weekend Activities, Paper presented at *2005 Transportation Association of Canada (TAC) conference*, Calgary, September 2005.
22. Lockwood, A.M., S. Srinivasan, and C.R. Bhat. (2005). An Exploratory Analysis of Weekend Activity Patterns in the San Francisco Bay Area. Presented at *the 84th Transportation Research Board Annual Meeting*, Washington DC.
23. Ettema, D., A. Borgers and H. Timmermans. (1995). Simulation model of activity scheduling behavior. *Transportation Research Record 1413*, pp. 1-11.
24. Chu, Y.L. (2005). Modeling Workers' Daily Non-work Activity Participation and Duration. Paper presented at *the 84th Transportation Research Board Annual Meeting*, Washington, DC, 2005.
25. Hamed, M.M. and Mannering, F.L. (1993) Modeling Travelers' Postwork Activity Involvement: Toward A New Methodology. Transportation Science, Vol. 27, No. 4, pp. 381-394.
26. Mannering, F.L. and Hamed, M.M. (1990) Occurrence, Frequency, and Duration of Commuters' Work-to-home Departure Delay. Transportation Research, Part B, Vol. 24, No. 2, pp. 99-109.
27. Nam, D. and Mannering, F. (2000) An Exploratory Hazard-based Analysis of Highway Incident Duration, Transportation Research, Part A, 34, pp. 85-102.
28. Paselk, T.A. and F.L. Mannering. (1994). Use of Duration Models for Predicting Vehicular Delay at A US/Canadian Border Crossing. *Transportation*, 21, pp. 249-270.
29. Stathopoulos, A. and Karlaftis, M.G. (2002). Modeling Duration of Urban Traffic Congestion. Journal of Transportation Engineering, Vol. 128, No. 6, pp. 587-590.
30. Hunt, J.D. and Atkins, D.M. (2004) Characteristics of Weekend Travel in Calgary. Transportation Revolutions: Proceedings of *the 39th Annual Conference of the Canadian Transportation Research Forum*, Calgary AB, Canada, May 2004, pp. 253-267.

31. MINITAB (2006). Minitab StatGuide. Available with the software.
32. Lewis, P.A.W. (1961) Distribution of the Anderson-Darling Statistic. *The Annals of Mathematical Statistics*, Vol. 32, No. 4, pp. 1118-1124.

Table 1. Percent difference for the weekday and weekend activity durations

	Travel etc.			Work			School			Shopping		
	weekday	weekend	Difference	weekday	weekend	Difference	weekday	Weekend	Difference	weekday	weekend	Difference
Mean	18.7	18.1	-3.5%	229.4	249.6	8.8%	175.3	148.7	-15.2%	44.9	50.3	11.9%
Median	15.0	15.0	0.0%	210.0	227.5	8.3%	150.0	120.0	-20.0%	30.0	35.0	16.7%
P75	25.0	20.0	-20.0%	300.0	369.0	23.0%	215.0	197.5	-8.1%	60.0	69.5	15.8%

	Sociality			Eating			Ent/Leis			Exercise		
	weekday	weekend	Difference	weekday	weekend	Difference	weekday	Weekend	Difference	weekday	weekend	Difference
Mean	82.0	128.3	56.5%	36.6	42.0	14.9%	125.4	143.9	14.7%	64.0	76.0	18.7%
Median	60.0	105.0	75.0%	30.0	30.0	0.0%	110.0	120.0	9.1%	60.0	60.0	0.0%
P75	110.0	165.0	50.0%	45.0	60.0	33.3%	180.0	210.0	16.7%	89.0	90.0	1.1%

	Religious, Civil etc.			Out of Town		
	weekday	weekend	Difference	weekday	weekend	Difference
Mean	63.9	112.8	76.5%	772.4	492.1	-36.3%
Median	20.0	105.0	425.0%	735.0	425.0	-42.2%
P75	79.3	130.0	64.0%	962.5	720.0	-25.2%

Table 2 Statistics for Log-Rank and Wilcoxon tests

Activity	Log-Rank		Wilcoxon	
	Chi-Square	p-value	Chi-square	p-value
A. Travel etc.	4.73	0.030	6.74	0.009
D. Work	3.25	0.071	0.65	0.418
E. School	3.47	0.062	5.70	0.017
F. Shopping	4.51	0.034	3.93	0.047
I. Sociality	35.88	0.00	45.58	0.00
J. Eating	46.89	0.00	47.96	0.00
K. Ent/Leis	30.41	0.00	24.99	0.00
M. Exercise	4.00	0.045	1.65	0.199
N. Religious, Civic, etc.	35.68	0.00	90.53	0.00
Z. Out of Town	3.90	0.048	7.89	0.005

Table 3 Goodness-of-fit tests for individual weekday and weekend activities

Distribution	Weekday/ weekend	Travel Related		Work		School		Shopping		Sociality		Eating		Ent/Leis		Exercise		Religious etc.		Out of town	
		AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR	AD	COR
Weibull	Weekday	63.02	0.97	6.28	0.99	3.71	0.99	6.21	0.97	**0.52**	**0.99**	76.09	0.96	5.35	0.99	1.81	0.98	10.83	0.89	1.26	0.97
	Weekend	103.15	0.96	1.434	0.99	0.87	0.99	9.58	0.98	0.68	1.00	84.05	0.96	5.27	1.00	4.01	0.96	3.38	0.97	**0.71**	**0.99**
Lognormal	Weekday	42.01	0.99	25.49	0.95	6.29	0.98	1.77	0.99	3.63	0.97	42.33	0.98	14.65	0.98	2.35	0.98	2.95	0.96	1.01	0.99
	Weekend	53.07	0.98	6.13	0.94	1.10	0.99	3.75	0.99	4.97	0.97	45.06	0.98	21.10	0.98	0.97	0.99	3.46	0.94	1.70	0.97
Exponential	Weekday	209.26	*	67.27	*	37.37	*	4.52	*	0.86	*	298.78	*	90.89	*	26.84	*	22.04	*	5.99	*
	Weekend	232.23	*	10.29	*	6.41	*	5.55	*	9.79	*	308.36	*	136.19	*	14.76	*	29.89	*	5.95	*
Loglogistic	Weekday	48.39	0.99	24.89	0.95	6.34	0.97	2.86	0.99	3.64	0.97	43.79	0.98	19.50	0.98	2.68	0.98	3.57	0.96	1.05	0.98
	Weekend	52.52	0.98	5.81	0.94	1.36	0.98	7.06	0.98	4.64	0.97	49.05	0.98	25.45	0.97	**0.95**	**0.99**	2.42	0.96	1.95	0.96
3-Parameter Weibull	Weekday	43.21	0.98	6.40	0.99	3.18	0.99	2.59	0.98	0.58	0.99	64.72	0.97	**4.83**	**0.99**	**1.35**	**0.99**	3.52	0.96	**0.99**	**0.99**
	Weekend	77.79	0.97	**0.84**	**1.00**	**0.51**	**1.00**	6.10	0.98	**0.67**	**1.00**	73.35	0.96	**4.77**	**1.00**	1.99	0.98	3.87	0.97	0.78	0.99
3-Parameter Lognormal	Weekday	**35.61**	**0.99**	**4.76**	**0.99**	**2.78**	**0.99**	**1.50**	**0.99**	0.99	0.99	**42.18**	**0.98**	7.23	0.99	1.50	0.99	**1.55**	**0.98**	1.01	0.99
	Weekend	**50.93**	**0.98**	1.22	0.99	0.66	0.99	**3.23**	**0.99**	0.95	0.99	**45.85**	**0.98**	9.33	0.99	0.97	0.99	1.89	0.98	0.80	0.99
2-Parameter Exponential	Weekday	140.34	*	64.33	*	31.13	*	10.30	*	2.86	*	252.10	*	86.41	*	14.55	*	13.57	*	1.53	*
	Weekend	167.58	*	10.02	*	3.58	*	3.34	*	9.00	*	266.61	*	130.67	*	6.77	*	26.20	*	3.57	*
3-Parameter Loglogistic	Weekday	46.18	0.99	6.83	0.98	3.27	0.98	2.86	0.99	1.95	0.98	44.17	0.98	16.28	0.98	2.05	0.98	1.74	0.98	1.05	0.98
	Weekend	52.23	0.98	2.42	0.98	1.15	0.99	7.07	0.98	2.16	0.99	50.14	0.98	21.72	0.98	0.98	0.99	**1.27**	**0.99**	1.14	0.98
Smallest Extreme Value	Weekday	455.84	0.67	90.13	0.89	60.49	0.85	71.22	0.68	36.03	0.71	262.12	0.79	131.90	0.86	12.10	0.88	24.29	0.51	1.77	0.94
	Weekend	519.70	0.66	21.25	0.89	13.46	0.84	131.05	0.78	52.77	0.84	261.01	0.79	172.64	0.87	22.19	0.75	29.04	0.83	5.03	0.92
Normal	Weekday	161.42	0.78	14.63	0.97	12.53	0.95	24.34	0.80	11.41	0.83	92.95	0.89	26.22	0.96	2.82	0.96	15.68	0.64	1.09	0.98
	Weekend	221.91	0.78	4.12	0.97	3.38	0.94	37.32	0.90	11.31	0.94	96.24	0.89	34.08	0.96	7.85	0.86	6.82	0.92	1.13	0.98
Logistic	Weekday	152.49	0.81	14.52	0.96	11.38	0.95	23.21	0.82	10.91	0.84	94.64	0.90	30.28	0.95	3.00	0.96	14.82	0.67	1.12	0.97
	Weekend	205.59	0.81	5.02	0.96	3.27	0.94	37.06	0.90	11.38	0.94	98.78	0.90	43.45	0.96	6.57	0.88	5.53	0.93	1.35	0.97

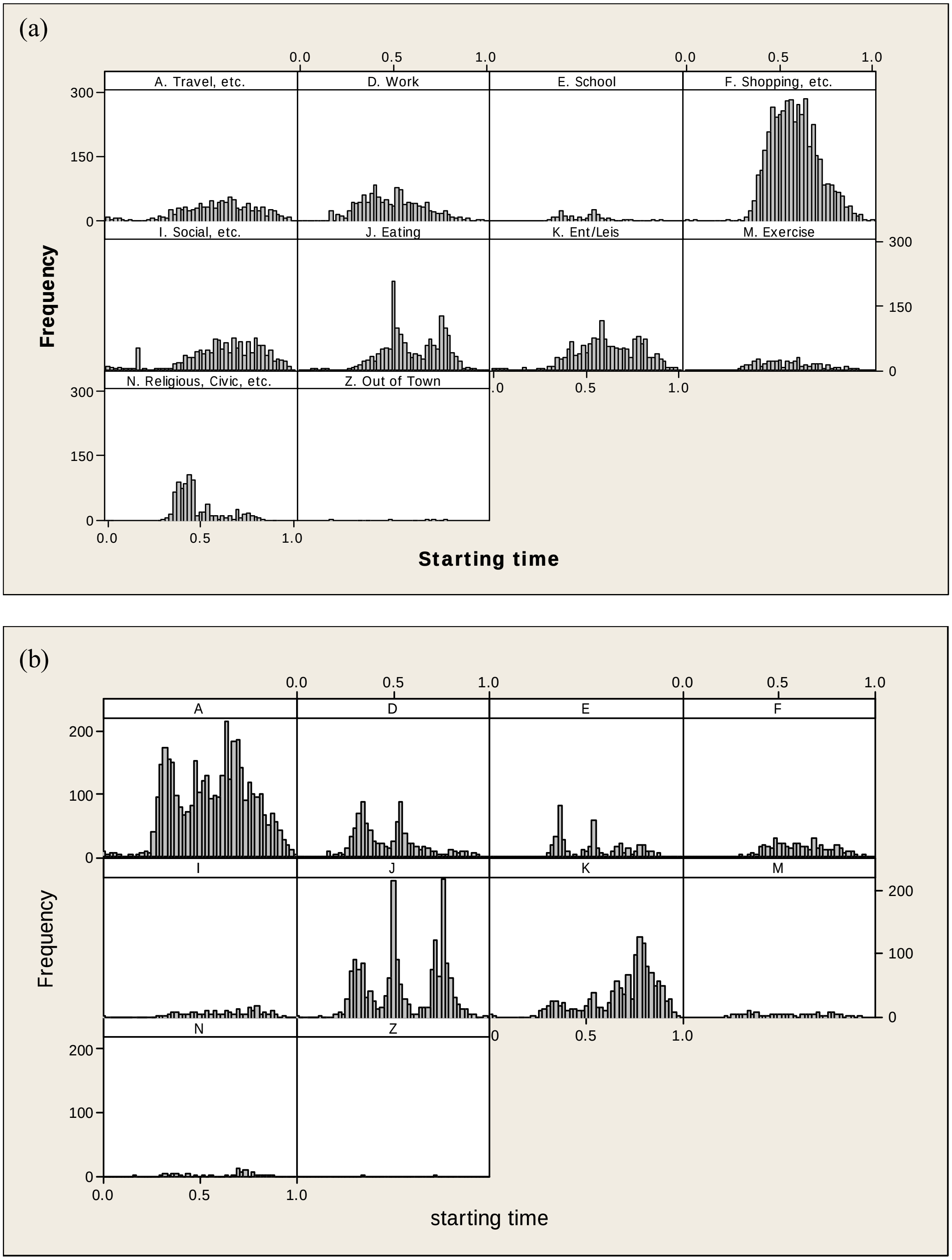
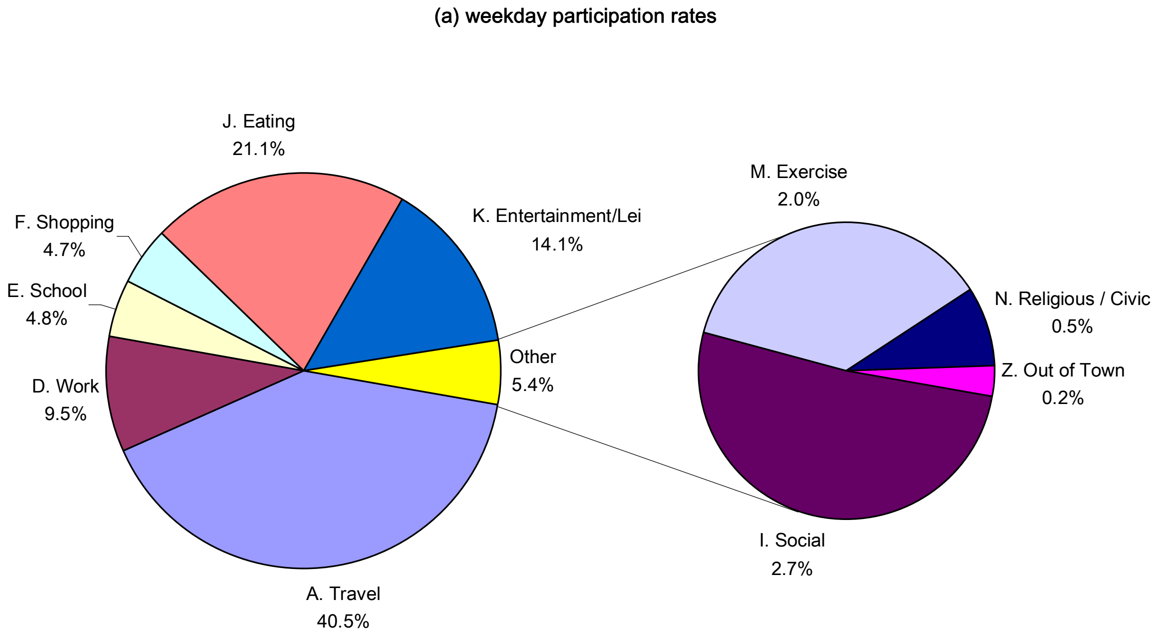


Figure 1 Comparison of household weekend (a) and weekday (b) activity patterns



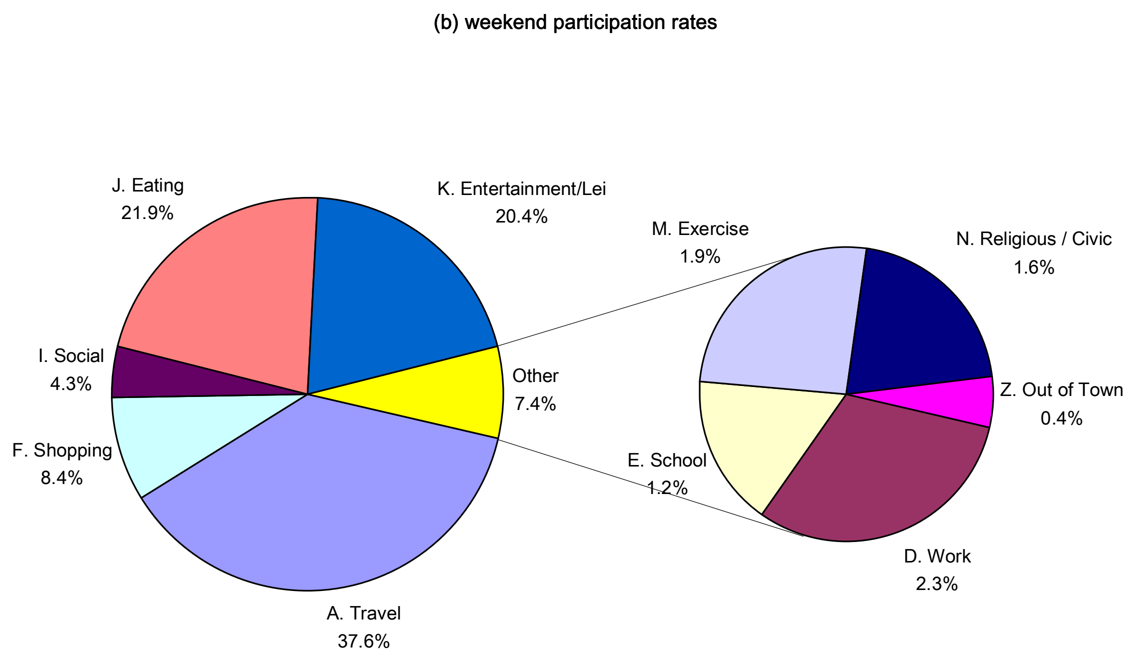


Figure 2 Calgary weekday and weekend activity participation rates

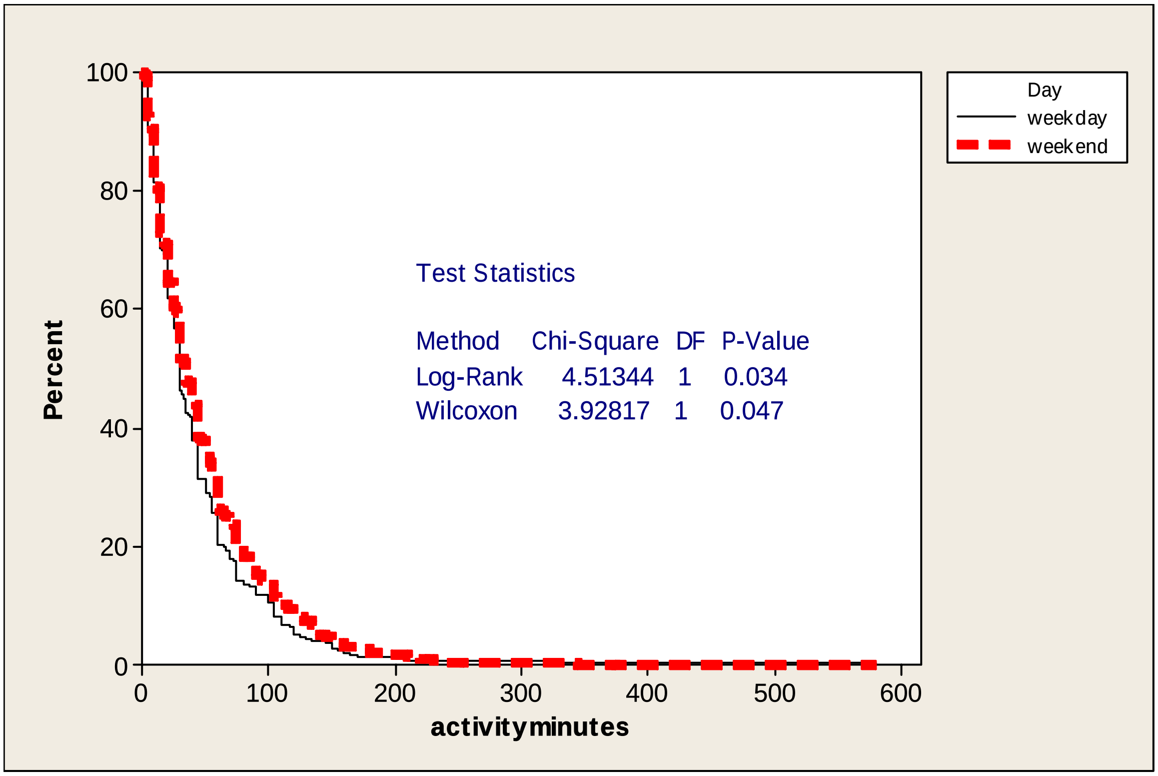


Figure 3 Kaplan-Meier test for weekday and weekend shopping activity

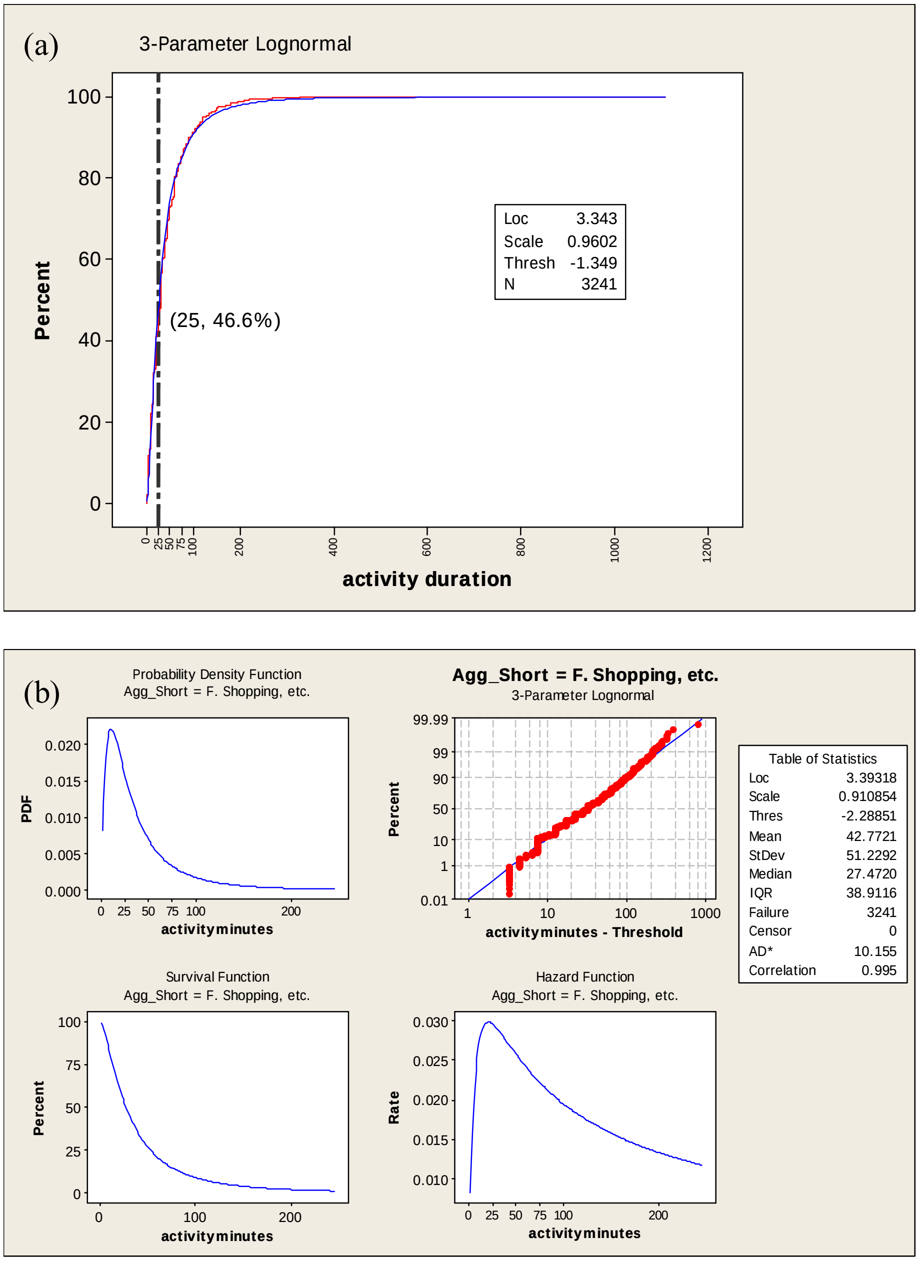


Figure 4(a) empirical cumulative distribution function (b) distribution overview plot for weekend shopping activities