multi-state non-homogeneous semi-markov model of daily activity type, timing and duration sequence

Tai-Yu Ma, Charles Raux, Eric Cornelis, Iragaël Joly

Multi-state non-homogeneous semi-Markov model of daily activity type, timing and duration sequence

Tai-Yu Ma, Charles Raux, Eric Cornelis, Iragaël Joly

Abstract

Understanding travelers' daily travel-activity pattern formation is an important issue for activity-based travel demand analysis. The activity pattern formation concerns not only complex interrelations between household members and individual's socio-demographic characteristics but also urban form and transport system settings. To investigate the effects of these attributes and the interrelationship between conducted activities, a multistate semi-Markov model is applied. The underlying assumption of the proposed model states that the state transition probability depends on its adjoining states. Based on the statistical tests of significance, it is affirmed that the duration of activity depends not only on its beginning time-of-day but also on the duration of travel/activity previously conducted. The empirical study based on Belgian Mobility survey is conducted to estimate individual's daily activity durations of different episodes and provides useful insight for the effects of socio-demographic characteristics, urban and transportation system settings on the activity pattern formation.

INTRODUCTION

The daily travel and activity patterns concern multidimensional facets encompassing activity type, duration of activities and the location of activities, etc. The formation of a sequence of travels and activities is general heuristical, based on individual schedule/reschedule process given an uncertain environment. The understanding of the complex travel-activity decision has important implications for the evaluation of transport policy and traffic demand management. Hence, it is interesting and also important to investigate individuals' travel-activity chaining behavior with respect to related household characteristics. The results of analysis can provide useful insights for travel and activity duration estimations and also for individuals' daily travel activity participation.

The early studies for activity patterns have focused on descriptive analysis of activity pattern characteristics and influence factors for its formation (*1*,*2*). The analysis framework emphasized on the concept that travel demand is derived from the need of activity participation. As numerous influence factors impose dynamical spatial-temporal constraints on activity pattern formation, it is generally difficult to investigate the relationships between these factors and the resulting travel-activity patterns. The progress in activity-based demand analysis has developed many modeling techniques trying to elucidate the process of activity and travel decisions. In general, these analytical approaches can be classified into: (a) sequential approaches, which decompose individuals' travel-activity choices into a sequential decision-making process and apply microsimulation techniques to entail the estimation of travel-activity patterns (see recent review in (*3*); (b) simultaneous approaches, which aim to estimate simultaneously the entire activity patterns by utilizing econometrics or mathematical programming methods (*4*,*5*,*6*,*7*). An alternative class of modeling approaches is based on hazard models for the empirical analysis of activity duration and sequencing (*8*,*9*,*10*). These studies focused especially on the estimation of activity duration hazard function with the presence of individual's heterogeneity. As for the interdependency of durations between multiple activity types, Bhat (*11*) proposed an outcome-specific proportional hazard model, which generalizes the estimation of duration hazard function with multi-entrance/exit states. Another activity hazard-based study, proposed by Srinivasan and Guo (*12*), encompassed the correlation of simultaneous duration process. This model utilized a mixed distribution combining baseline hazard function with a random error term to capture the unobserved correlation between two duration processes. This study provided empirical evidence of dependency on the durations of trips and activities. However, the considerations of the influence of activity sojourn time in hazard function specification need to be further investigated.

The time dependency effect plays an important role in the estimation of activity duration since it reflects the transition probability after certain duration from current activity to another. Recently, some research effort has been made to investigate the relationship of timing and duration of activity. Pendyala and Bhat (*13*) applied discrete-continuous simultaneous equation model to investigate the casual structure of activity timing and duration. They found that the activity timing and duration are closely related for non-commuters but loosely related for commuters. Popkowski Leszczyc and Timmermans (*14*) utilized conditional and unconditional parametric competing risk models to estimate activity duration and its relationship with socio-demographic covariates. The estimation results showed that the activity duration depends on the type and duration of the activity previously conducted. Parallel to the competing risk models for the estimation of timing and duration of activity, Lee et al. (*15*) proposed a generalized multi-state hidden Markov extension model to analyze the duration of type-specific activities. Although the proposed model provided a flexible framework to capture the misclassification effect of observed activities, it did not incorporate the dependency effect of the duration of activity previously conducted. Recently, the advance in survival analysis has developed multi-state models to capture simultaneously several entrances/exits states over the duration process (see review in (*16*)). The multi-state models analyze event history data of individuals, focusing especially on the occurrence time, duration of type-specific events following a Markov process. In general, the multi-state models consider event occurrence history based on Markov-type stochastic process, i.e. step transition probability depends only on the adjoining states. At each time instant, an individual occupied in one state possibly transfers to another state with respect to time dependent/independent transition rates. As the activity duration depends on its type, timing and sojourn time prior to current state, it is interesting to investigate the influence of time dependency effects on the adjoining travel/activity states for individual's activity chaining behavior.

In this study, a multi-state semi-Markov model (*17*,*18*) is applied to estimate the daily activity patterns by encompassing the interdependency of sequential activity types, timing and duration. The daily travel/activity duration sequence is assumed to result from semi-Markov process for which the observed duration is independent continuous random variable with distribution depending on its adjoining states. Based on the assumption of semi-Markov model, the state transition probability is estimated with Cox regression model (*19*). We associate the socio-demographic, urban form and accessibility of public transport covariates with the specification of transition hazard function to investigate the effects on the travel/activity duration. The transition hazard rates between different states are estimated with respect to cause-specific covariates. These estimated coefficients provide also useful insight to investigate the influence of socio-demographic characteristics, urban form and transportation system settings on the activity chaining behavior.

The remainder of this paper is organized as follows. First, a multi-state semi-Markov model is proposed to analyze individual's daily activity pattern with respect to activity timing, sequencing and duration. The underlying assumption of proposed model based on semi-Markov process is firstly discussed. It follows multi-state dependency structure of daily activity pattern formation. For the state transition probability estimation, three Cox regression models are specified to incorporate and compare the timing and sequencing effects on travel/activity durations. The summarized statistics of travel survey data and observed numbers of travels/activities conducted in different episodes are reported. We discuss first the model-fit statistics and the assessment of proportionality assumption of the Cox regression model for each of state transitions in the travel-activity chaining process. The estimation result of baseline hazard for each of non-travel activity purposes is discussed and compared over each of episodes of interest. We investigate the effect of related covariates on the travel/activity duration over episodes and discuss in detail these effects. Finally, the conclusion is drawn and the future extension is discussed.

MULTI-STATE NON-HOMOGENEOUS SEMI-MARKOV MODEL

Consider a travel-activity sequence of an individual observed over a whole day for which the types of activity, timing and duration are of interest. This sequence is derived from individual's daily activity scheduling and rescheduling decision under uncertain environment. The ordered travel/activity duration sequence represents individual's daily time-use pattern. This pattern concerns subsequent activity choices influenced by related temporal and spatial factors. The temporal factor expresses that the time being in an activity is related to its occurrence of time-of-day, the time elapsed since entering occupied state and the sojourn time on travel/activitie previously conducted. The spatial factor describes that activity destination choice is related to the opportunities offered in an area and the location of activity impacts in turn the available time to participate the activity. To derive this activity pattern in terms of timing, sequence and duration of activity, some underlying assumptions are stated as follows:

(i) Individual's daily activity pattern is represented as time dependent evolution of activity participation states. The state space is finite and assumed to be identical for all individuals. The evolution of activity states follows a semi-Markov stochastic process, i.e. the time-dependent state transition probability depends only on its adjoining state.

(ii) The state-to-state transition is considered as an alternate of non-travel activity state and travel state. The episode of one state represents the duration being in that state in which the conducted activity is assumed homogeneous without state transitions within an episode.

(iii) The duration of activity is an independent random variable depending on its adjoining state, the sojourn times being in an occupied state and related covariates.

Model formulation

Based on the above assumption, individual's daily activity pattern observed throughout 24 hours can be expressed as the evolution of states in time, $\{a(t), 0 \le t \le \overline{T}\}$ with $\overline{T}$ being the ending time of observation period, $a(t)$ the travel/activity state and A the finite state space of one transition. This sequence of states, corresponding to individual's travel/activity occurrences over time, is denoted as (see Fig. 1):

$$q = (a_i, t_i) \quad i \in \{0,1,2,...,n\} \tag{1}$$

where

a_i: state of the i-th travel/activity, $\forall a_i \in A$

t_i: beginning (transition) time of the i^{th} travel/activity, with $0 \le t_0 < t_1 < ... < t_n \le \overline{T}$.

The initial state of the activity pattern is set as home-based activity at the beginning of a day. As the initial time of the beginning activity and the ending time of the last activity are unknown within the 24 hour observation period, the duration of initial/terminal activities is censored. We restrict ourselves on the analysis of uncensored observations only. Note that although the occurrence time of activity is continuous, the available precision of data in practice is rounded in minute. The estimation of transition probability based on the empirical survey data should consider the ties of travel/activity duration (described later).

Based on the observed entering/exit time of travel/activity in the sequence, the duration of the n-1th travel/activity τ_n is defined as $t_n - t_{n-1}$. An individual's occupied state history on the process until time t is denoted as:

$$H_t = \{a(s), 0 \le s \le t\} \tag{2}$$

To estimate the time-dependent transition probability, the hazard model is applied to this end. The hazard rate represents the instantaneous changing rate of failure probability at time t, given that the components in observation are survival at time t. As the probability distribution of activity duration is assumed independent following the semi-Markov process, the state transition probability can be estimated separately based on usual hazard model. The interdependency between the sequential states visited is determined by incorporating relevant time dependent/independent covariates. The one-step transition hazard $\lambda_{ij}(\tau(t))$ from occupied state i to next state j, j=1,…,m, at time t is function of sojourn time $\tau(t)$ at state i until t^- since entering occupied state i. We call the sojourn time since entering current state as *renewal time*. For simplification, $\tau(t)$ is denoted as τ hereafter. The corresponding cause-specific (type-specific) transition hazard function is defined as:

$$\lambda_{ij}(\tau) = \lim_{h \to 0} \frac{P\left[\tau \leq \mathrm{T} < \tau + h, a(\tau + h) = j \,\middle|\, a(\tau) = i \right]}{h} \tag{3}$$

where T is a continuous random variable representing travel/activity duration, $a(t^-)$ is the state at time t^-. The transition rate $\lambda_{ij}(\tau)$ represents the changing rate of conditional probability of transition from state i to state j at renewal time τ. In the proposed model, it is assumed that only one travel/activity type can occur at a time. To estimate the transition hazard, parametric, semi-parametric or non-parametric approaches can be applied. As there is no prior information for the functional form of hazard function, the semi-parametric Cox regression model is preferred since it does not need to specify the underlying probability distribution and can incorporate related covariates of interest. The Cox regression model assumes that different classes of individuals have similar hazard profile of travel/activity duration. The relative risk of stopping a travel/activity episode is proportional to common baseline hazard, determined by the value of related covariates. Different methods have been proposed in the literature to assess the proportional hazard assumption (described later). In the above cause-specific hazard function, the duration of activity depends only on the sojourn time being in occupied state. However, it is important to incorporate the occurrence time of activity and the dependency of adjoining travels/activities in the hazard function specification. To this end, one can specify the baseline hazard as function of renewal time and chronological time. However, it is more difficult to estimate since it needs to specify the kernel estimate for the regression parameter estimation (*17*, *27*). An alternate way consists in incorporating these terms in the covariates and utilizes traditional parameter estimation method of Cox regression model. We specify three models to assess the effect of occurrence time of activity and the interdependency between adjoining states.

Model 1: including related covariates in terms of individual's socio-demographic characteristics, urban form and accessibility to public transportation system. The included variables are selected based on stepwise regression method.
Model 2: extending model 1 by incorporating entering time of occupied activity.
Model 3: extending model 2 with the duration of travel/activity conducted in previous episode.

The estimation of parameters is based on the maximum partial likelihood estimate. The Log-likelihood statistics of the above models are compared in order to assess the fitness of models. The best-fit model is used as final model for the statistical inference on the travel/activity time-use pattern.

For one-step state transition hazard estimation, the Cox model is specified with respect to related covariates. The transition hazard for covariates $\mathbf{X}_{ij}$ from state i to j is defined as:

$$\lambda_{ij}(\tau; \mathbf{X}_{ij}) = \lambda_{0,ij}(\tau) \exp(\mathbf{X}_{ij}^{'} \boldsymbol{\beta}), \quad \forall i \neq j \tag{4}$$

where
$\lambda_{0,ij}(\tau)$: unspecific baseline hazard function with respect to transition (i, j)
$\mathbf{X}_{ij}$: column vector of transition-specific time-independent covariates for (i, j)
$\boldsymbol{\beta}$: column vector of regression coefficients.

Note that the above usual exponential function specifies the relative risks with respect to adequate covariates. It assumes that all adequate covariates are incorporated and the covariates have the multiplicative effects on hazard function.

As for the explanation of regression coefficients in Cox regression model, it represents the relative risk on the basis of baseline hazard. If the regression coefficient is positive (negative), the relative risk is increasing (decreasing), i.e. activity duration is decreasing (increasing). Note that the baseline hazard function is transition-specific and common for all individuals under observation. The observed heterogeneity is taken into account with respect to the covariates. For the unobserved heterogeneity, it is assumed to have random effect on hazard function. However, one can incorporate a parametric term in the hazard function specification to consider the unobserved heterogeneity effects (*9*, *10*).
The Cox partial log-likelihood for transition (i, j) can be written as (*17*):

$$\log L(\boldsymbol{\beta}) = \sum_{l}\sum_{ij}\int_{0}^{\bar{T}}\left\{\mathbf{X}'_{ijl}\boldsymbol{\beta} - \log[\sum_{m} Y_{im}(\tau)\exp(\mathbf{X}'_{ijm}\boldsymbol{\beta})]\right\}dN_{ijl}(\tau) \tag{5}$$

where
$N_{ijl}(\tau)$: observed number of transition from state i to j for individual l within time interval $[t_{il}, t_{il} + \tau]$. t_{il} is the entering time of current state i of individual l.
$Y_{im}(\tau)$: indicator being 1 if individual m is at state i under risk within $[t_{il}, t_{il} + \tau]$; 0 otherwise.
Note that $N_{ijl}(\tau)$ is the usual notation of counting process for survival analysis. It represents the right counting process of observed number of transition from i to j of individual l on $(0, \tau]$. The notation $dN(\tau)$ in continuous time defined as $N(\tau^{-} + d\tau) - N(\tau^{-})$, representing the number of transition within the renewal time interval $[\tau, \tau + d\tau)$. As for the parameter estimation, the maximum likelihood estimate for $\boldsymbol{\beta}$ can be derived by applying the usual Newton-Raphson method (*20*). Note that if there are tied lifetime (duration) data, i.e. the same observed travel/activity duration for one episode, some approximated approaches may be used to handle this issue. The exact approximation of Cox partial likelihood with tied duration data considers the duration distribution is continuous and calculates the partial likelihood approximate with respect to all possible ordering of tied durations (*20*). This method requires a considerable compute time and computer memory if the number of tied lifetime data is large. In practice, if the proportion of tied data is small, Breslow method (*21*) provides close results with respect to the exact approximation method. However, if the number of tied data is large, Efron method (*22*) can be applied. The presence of tied duration data is very common in travel survey since the departure/arrival time reported by individuals is usually rounded by 5 minutes. Hence, this issue needs to be handled by adequate approximation method for empirical study.

As the baseline hazard function in Cox regression model is unspecified, the empirical approximate is needed to estimate the hazard function and the survival function. Based on Nelson-Aalen estimator, the usual empirical cumulative hazards function $\hat{\Lambda}_{ij}(\tau)$ can be estimated as:

$$\hat{\Lambda}_{ij}(\tau) = \sum_{l}\int_{0}^{\bar{T}}\frac{dN_{ijl}(\tau)}{\sum_{m} Y_{im}(\tau)\exp(\mathbf{X}'_{ijm}\boldsymbol{\beta})} \tag{6}$$

with $N_{ijl}(\tau)$ and $Y_{im}(\tau)$ being the same notation as in the equation (5).

The transition hazard estimate with respect to the covariates $\mathbf{X}_{ij}$ is then written as:

$$d\hat{\Lambda}_{ij}(\tau) = 1 - [1 - d\hat{\Lambda}_{0,ij}(\tau)]^{\exp(\mathbf{X}_{ij}^{'}\boldsymbol{\beta})} \tag{7}$$

where $\hat{\Lambda}_{0,ij}(\tau)$ is the cumulative baseline hazard estimate with respect to transition (i, j).

With $\mathbf{X}_{ij} = \mathbf{0}$, the transition-specific baseline hazard estimate can be obtained by (6) and (7).

The survival function estimate $\hat{S}_{ij}(\tau)$ of transition from *i* to *j* at renewal time τ can be calculated as:

$$\hat{S}_{ij}(\tau) = \exp(-\hat{\Lambda}_{ij}(\tau)) \tag{8}$$

In individual's daily activity sequence, the travel/activity transition may be distinguished as non-travel state to travel one and the inverse case. The first type of transition describes one possible state transition at the end of non-travel activity episode. However, the second type corresponds to the competing risk over several distinct non-travel activities. At each time instant, travel state may terminate at one of the competing non-travel activities, which guarantees the sum of failure probabilities over competing causes as 1. Mathematically, the transition hazard at state *i* can then be written as

$$\lambda_i(\tau;\mathbf{X}) = \sum_j \lambda_{ij}(\tau;\mathbf{X}_{ij}) \tag{9}$$

where *j* represent one of competing non-travel activity choice at the end of a trip at renewal time τ.

The assessment of proportionality assumption

The Cox model with time-independent covariates describes the fixed proportional effects of covariates on the duration of travel/activity. If the proportionality assumption is not valid, the statistical inference will be biased. Numerous methods have been proposed for the test of proportionality (*20*, *23*, *24*). A usual technique for model-fit test is based on the distribution of residuals, i.e. the differences between observed and predicted value of the model. The basic concept is that the distribution of the residuals should be centered at zero if the model is correctly specified. Lin et al. (*25*,*26*) proposed a graphic and numerical test method based on the cumulative sum of residuals over covariates. The proposed graphic test method compares observed cumulative sums of the residuals and a large number of simulated realizations based on zero-mean Gaussian process. If the Cox model is correctly specified, the observed cumulative sums of residuals should centered at zero and the simulated realizations should presents randomly frustrations around the observed cumulative sums of residuals. Moreover, the cumulative residual plot suggests appropriate functional form for a covariate with non-proportional effects. For the numerical test, the estimated p-value for the supremum test over covariates can be used as an criteria to assess the proportionality assumption of the Cox model. The assessment of the proportionality assumption for the empirical study will be conducted based on Lin's approach.

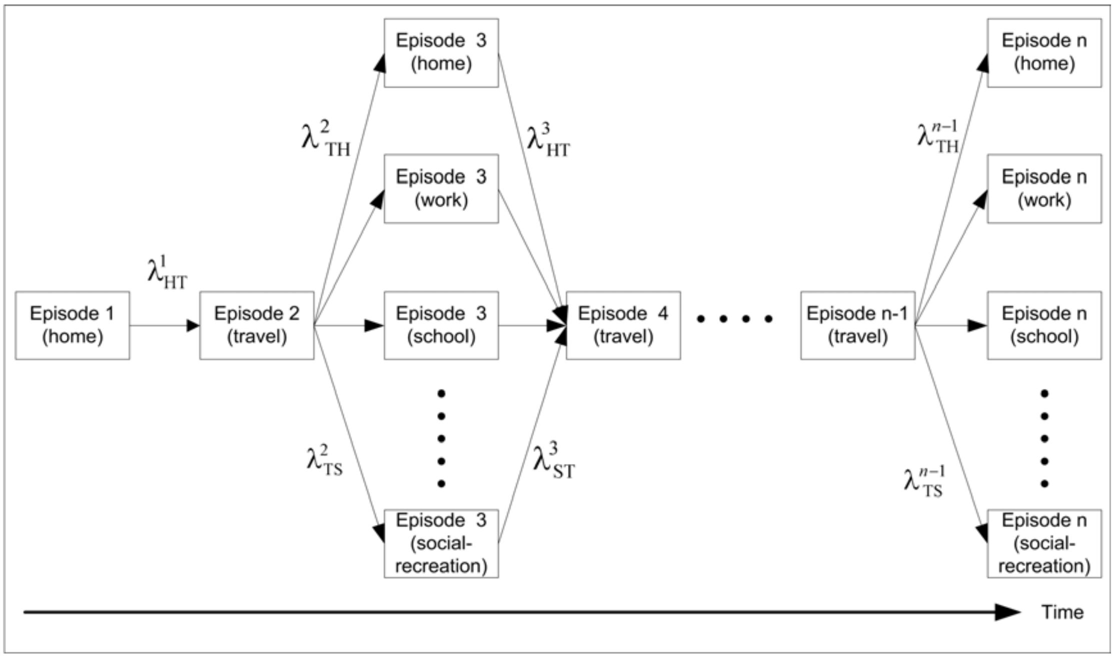


Figure 1 Non-homogeneous semi-Markov process for travel-activity pattern formation

DATA DESCRIPTION

The data utilized is based on the Belgian Mobility Survey collected in 1998-1999 for investigation of household members' mobility behavior. 7025 observations of individual's daily travel-activity patterns were collected. As our interest resides on the investigation of travel-activity time use pattern, the focus is on the sequence of travels and out-of-home activities realized on a day. Hence, samples without any trip occurred are not included in the data set of analysis. After comparing the activity participation on weekday and weekend, we found that individual's activity choice patterns are quite different between them. The former is work/school oriented. However, the latter is shopping and social-recreation dominated. This study focuses only on travel-activity pattern over workday. One can, however, extend similar analysis for time use data conducted on weekend. Based on previous studies (*10*), some individual's (gender, age and profession status) and household socio-demographic characteristics (household type, and presence of children) are included in the model specification. Besides, spatial and transport mode related variables (household location, transport mode availability and proximity to public transport) are also included in the model specification. The summary statistics of these covariates is listed in Table 1.

As for the travel-activity sequence data, individuals are asked to report his/her sequence of trips and its related trip motivation on the previous day. For each reported travel/activity episode, trip departure/arrival time information is utilized to calculate trip duration. Original trip motivation is distinguished into 12 types for which a regroupment in 6 types is undertaken to facilitate the analysis (see Table 2). For simplification, we assume that the activity within two consecutive trip episodes is homogeneous. As a result, the activity duration is calculated based on the arrival/departure time of its prior/posterior trip. As the model estimation requires the full information for covariates, a complete data set is obtained by eliminating observations containing missing or incorrect data. The resulted data set contains 2849 observations and its related descriptive statistics is listed in Table1. For activity chaining behavior, the residual observations at each activity episode is shown in Table2. It is clear that home-work-home constructs the main travel-activity pattern on weekdays. The proportion of individuals who conducted less than 8 non-travel activities in a day is 84%.

Table 1 Summary statistics of covariates in terms of individual's socio-demographic, urban form and transport accessibility (n=2849)

Variable	Definition	Means	Std. dev.
Socio-demographic characteristics			
Gender	Gender (1 if male, 0 female)	0.51	0.50
Age15	1 if the age of the individual is less than 15 years, 0 otherwise	0.16	0.37
Age15_25	1 if the age of the individual is on $[15, 25)$ years, 0 otherwise	0.14	0.35
Age25_55	1 if the age of the individual is on $[25, 55)$ years , 0 otherwise	0.49	0.50
Age55_65	1 if the age of the individual is on $[55, 65)$ years, 0 otherwise	0.10	0.30
Age65	1 if the age of the individual is equal or greater than 65 years, 0 otherwise	0.11	0.31
Prof	Employment status (1 have a job, 0 otherwise)	0.47	0.50
H_type	Household type (1 if couple, 0 if single)	0.17	0.38
N_children	Number of young children less than 16 years of age in the household	0.90	1.25
Urban form and transport availability characteristics			
H_location	Location of household (1 center city/village, 0 otherwise)	0.61	0.49
N_car	Number of cars in the household	1.31	0.78
PT_proximity	Proximity to public transport (1 distance less than 250m, 2 250m$\leq$ distance$<$ 500, 3 500m$\leq$ distance$<$ 1km, 4 1km$\leq$ distance<2km, 5 2km$\leq$ distance$<$ 5km, 6 more than 5km)	1.96	1.12

Table 2 Number of individuals observed in sequential non-travel activity episodes

Type of activity	AEP1	AEP2	AEP3	AEP4	AEP5	AEP6	AEP7	AEP8	AEP9	AEP10	AEP11	AEP12	AEP13
Home	2849	40	1707	430	737	257	269	109	99	34	32	9	14
Work	-	927	260	239	108	69	38	20	17	6	4	3	1
School	-	645	28	103	8	12	5	5	0	0	0	1	0
Shopping	-	484	242	195	99	90	28	29	9	7	2	3	2
Personal business	-	442	237	259	132	109	59	52	18	16	5	6	1
Social-recreation	-	304	231	305	124	121	65	43	17	15	3	8	2
Other	-	7	8	3	1	3	0	0	1	1	0	0	0
Total	2849	2849	2713	1534	1209	661	464	258	161	79	46	30	20

Remark: 1. AEP: non-travel activity episode

2. The original trip purpose "go home" is regrouped in *Home*; The purposes "go to work" and "visit for work" are regrouped in *Work* activity, The purpose "go to school" is regrouped in *School*, the purposes "buy something/shopping" is regrouped in *Shopping*, The purposes "deposing/looking for somebody", "eating", and "personal business (bank, doctor etc.)" are regrouped in *Personal business*, The purposes "visiting families or friends", "take a walk", and "leisure, sport and culture activities" are regrouped in *Social-recreation*.

In the following, we estimate activity duration hazard functions and related survival functions for state transition based on Cox regression model. The estimation process consists in establishing a class of models with covariates of interests and then selecting a best-fit model based on model fit statistics. The stepwise regression method is utilized to select predictive covariates for which all candidates are included at the beginning, then tested one-by-one. The elimination is conducted for non-significant variables according to its statistical significance. Moreover, as the number of travel-activity episodes is large, we restrict ourselves on the analysis of the first four travel and non-travel activity episodes, each accounting for at least 58% observations of total data set.

ESTIMATION RESULTS

The parameter estimation is conducted by the Proc Phreg in SAS. As there are some tied duration data, the Efron method is used to compute the partial likelihood. We discuss first the model fit statistics for different model specifications and then assess the proportionality assumption for the state transition hazard estimates. Based on the Nelson-Aalen estimator, the profile of baseline hazard of type-specific activities is plotted with respect only to non-travel activity conducted in 4 episodes of study. We discuss these baseline hazard profiles and compare the difference between the episodes. Finally, the parameter estimates of transition-specific covariates across different episodes are discussed.

Model fit statistics and proportionality assumption test

The overall goodness-of-fit statistics and the likelihood ratio test for the comparison of different models are shown at Table 3. The log-likelihood value shows that the model 3 incorporating logarithm of the entering time of current travel/activity and logarithm of the duration of travel/activity conducted in previous episode performs best over the other two models in most transitions over the episodes. Hence the model 3 is preferred as the final model. The results show that the duration of travel/activity closely related to its type, beginning time-of-day and the duration of prior activity conducted, accordant with the empirical estimation results of activity duration of Popkowski Leszczyc and Timmermans (*14*). The likelihood ratio tests for the null hypothesis that all coefficients of covariates are zeros is rejected at 0.01 statistical significant level for the transitions of study across these episodes.

As for the proportionality assumption test, the model fit test based on the cumulative sums of residuals are conducted for included covariates over the transitions. The misspecification of Cox regression model is verified at 0.1 statistical significance based on supremum test (*25*). The estimated results indicate that the proportionality assumption of the final model is not verified only for two covariates among 172 ones in 8 episodes of interests (EP2-EP9). It accounts only 1.16% of covariates violating of aforementioned assumption. The covariates are "logarithm of entering time of activity" for home-trip transition in episode 7 (p-value is 0.059) and for school-trip transitions in episode 9 (p-value is 0.003). It indicates that these two covariates have not proportional effects on corresponding baseline hazard.

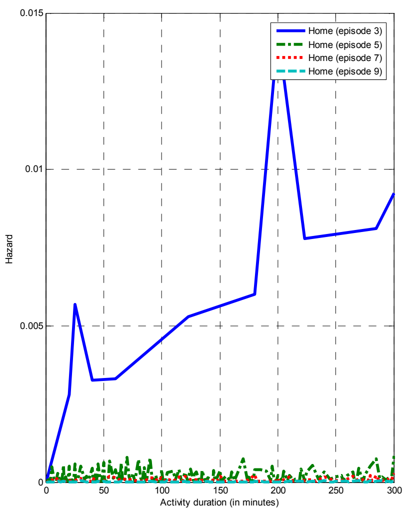
To summarize, the proportionality assumption of incorporated covariates is generally verified with only few exceptions. For the covariates with non-proportional effect, more adequate nonlinear functional form can be estimated with more elaborated methods (*26*).

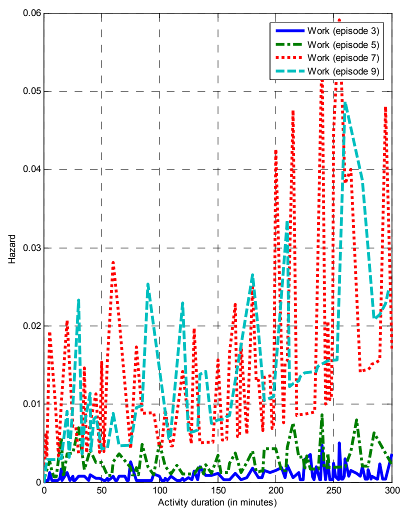
Baseline hazard estimation

As the number of travel/activity episode is large, we focus only on the baseline hazard estimation for non-travel activity episodes. The baseline hazard estimates for each of activity purposes conducted in different episodes are shown in the figure 2. To compare the baseline hazard conducted in different episodes and facilitate the lecture, the baseline hazard is plotted over the duration of activity from 0 to 300 minutes. The profile of baseline hazard provides the variation of instantaneous rate of terminating an activity for individuals in risk. The comparison of baseline hazard is conducted with respect to the first four non-travel episodes. The estimates of baseline hazard for each of activity purposes provide useful information to investigate the difference of temporal rhythm over different types of activities. Note that the median of the starting time of the four non-travel activity episodes (EP3, EP5, EP7 and EP9) is 8:30, 14:07, 15:00 and 16:45, respectively. It represents four temporal rhythms of individual's time use pattern starting in different periods of day.

For home activity, the baseline hazard on the EP3 exhibits quite different profile than the other 3 episodes. As the episode 3 is concerned, the result indicates that people staying at home in the morning tends to engage other out-of-home activities. The baseline hazard on the EP3 increases generally with time with the first spike at the duration of 25 minutes and the second at 200 minutes. This implies that people start his second out-of-home trip with these two temporal rhythms. On the other hand, the baseline hazard profile on the EP5 exhibits numerous spikes almost every 5 minutes within the first 120 minutes with a decreasing frequency of occurrence over time, suggesting that generally random home duration patterns in this afternoon period. For the EP7 and EP9, the baseline hazard is relative lower with respect to EP3 and EP5, implying that later the time of return home is, lower is the probability of engaging additional out-of-home activity. For work activity, it is reasonable to find the inverse temporal rhythm as the home activity. The result shows that the baseline hazard on the later episodes (EP7 and EP9) exhibits frequently sharp points every 10-15 minutes. It indicates the variable duration pattern of work during the afternoon period. For the school activity, the baseline hazards on the EP3 reveal that a higher activity terminating hazard at 25, 75, 135 and 200 minutes. For the episode 5, the result indicates a monotone increasing tendency on the baseline hazard. For the episode 7, the baseline hazard increases significantly as the duration of activity more than 2 hours. For shopping activity, the result indicates quite variable shopping duration pattern conducted in different episodes. In the EP3 and EP5, the baseline hazard profile shows a similar variation pattern of shopping duration within 2 hours, but a longer duration pattern of more than 3 hours for EP3. As for the EP7 and EP9, the variation of duration is in the range of 70 minutes. As the time elapse, the baseline hazard increases, i.e. shopping activity tends to be terminated. For personal business activity, the baseline hazard profile of the EP3 indicates that numerous spikes are spread over 0 and 280 minutes with a higher intensity within the first 80 minutes. On the other hand, for the EP5, EP7 and EP9, it reveals that a variation of short duration pattern less than 80 minutes and a longer duration pattern at 120, 160 and 210 minutes. For social-recreation activity, the baseline hazard profile for these episodes indicates a general variation of duration ranging from 0 to 300 minutes. The result shows a general similar temporal pattern over the morning and the afternoon periods.

To summarize, the baseline hazard profiles for each of activity purposes manifest different time use pattern over episodes, indicating different temporal rhythm of individual's activity participation. The comparison of the baseline hazard profiles over the activities showed a significant difference, revealing that activity duration depends on its type, starting time-of-day and the elapsed time since staring occupied activity. It is also shown that the baseline hazard function is not smooth and manifests an irregular pattern over time. It suggests that the parametric model may not be appropriate for the baseline hazard estimation. This result was also found by previous empirical study implemented by Bhat (*9*). He utilized a proportional hazard model for the estimation of shopping duration during the return home trip by incorporating the non-parametric heterogeneity distribution. As for the non-parametric baseline hazard estimate of shopping duration, it is interesting to find the similar behavior pattern of short shopping duration in EP9 (period starting the returning home trip after work) as Bhat's study. Moreover, our study revels that the shopping duration pattern manifests a range of variation over different period of time, i.e. later the shopping activity begins, smaller the range of variation of duration conducted tends to be.





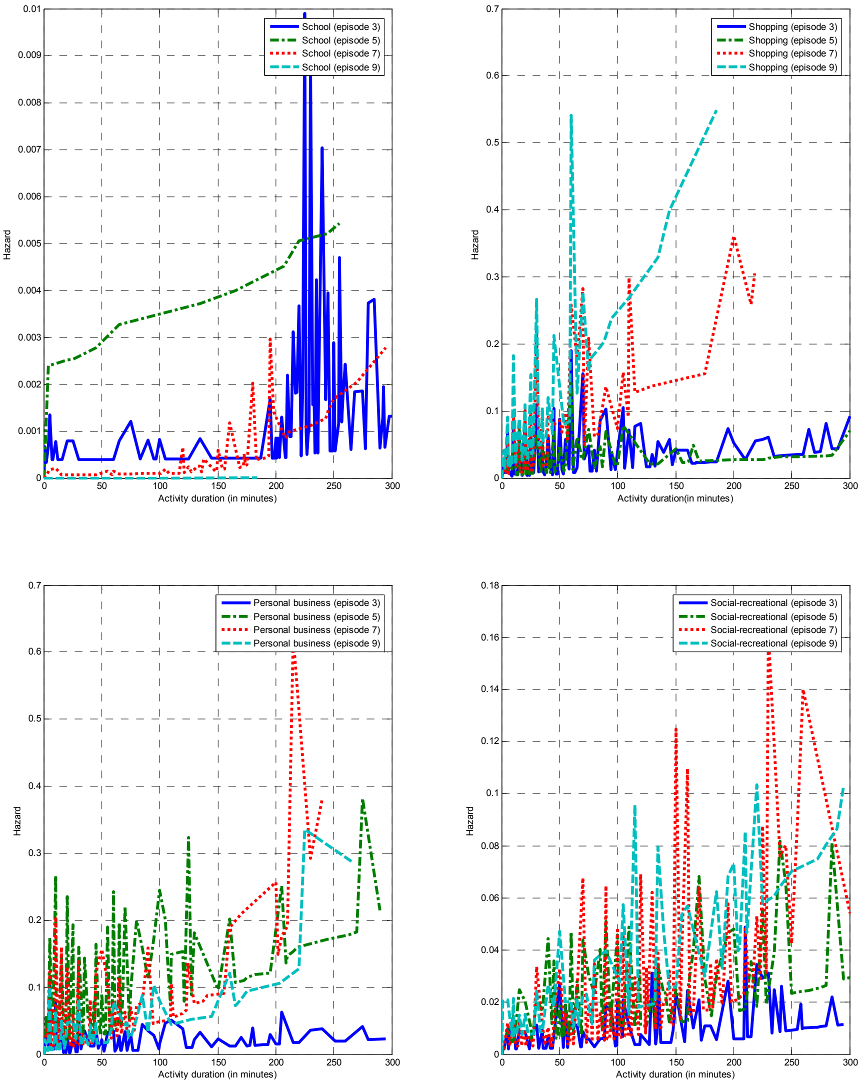


Figure 2 Baseline hazard estimates for activities conducted in episode 3, 5, 7 and 9

Covariate effects

The estimated effects of covariates on the travel/activity duration hazard conducted in different episodes are shown in Table 5. The covariates are distinguished as socio-demographic, urban form, transport accessibility and state-dependent covariates, i.e. entering time of occupied activity and duration of travel/activity previously conducted. Note that the duration process of travel episodes and activity episodes are in general different, which means that individuals tend to shorten his/her travel time in order to have longer non-travel activity duration. We discuss separately the effects of covariates on the two categories of activities to highlight their differences. Note that the effect of covariates depends on the starting time of day and varies over different periods in a day for activity participation.

Effects of covariates for non-travel activity episodes

In this section, we discuss firstly the effects of covariates on non-travel activity duration conducted in different episodes. The effects of covariates are discussed with respect to 4 aforementioned categories of covariates for each of activity purposes. It is of interest to investigate the time-of-day effect and the dependency between the duration of travel and non-travel activity in adjoining episodes and compare the effects of covariates in different episodes of activities.

First non-travel activity episode (EP3)

The first non-travel activity occurs near 8:30. The parameter estimates are significant for most socio-demographic covariates (Table 5). For home activity, the results indicate people of age more than 65 years stay shorter in home for the early morning episode. For work activity, it shows that men have longer duration. However, people perform shorter duration of work if they have children. For school activity, it is interesting to find that women, people of age more than 65 years, workers and singles perform shorter duration. For shopping activity, people of age between 25 to 55 years perform shorter duration but couple and people with the presence of children conduct longer duration of shopping. The results may results from the disavailability of shopping time in the morning for workers, generally in the group of age 25 to 55 years, and the additional maintenance need for couple and people with children. For personal business starting in the morning period, the results imply that people of age between 15 to 55 years, workers and people with the presence of children conduct shorter duration. As for social-recreation activity, the results suggest that people of age more than 65 years and couple spend more time on social-recreation activity. As for urban form covariate, people living in city center have shorter duration for home, work, personal business and social-recreation activities in the morning. It implies that individuals living in city center tend to conduct another out-of-home activities due to better accessibility for related activities in urban area. As for transport accessibility covariates, the parameter estimates indicate people with the accessibility of car conduct shorter duration of work and personal business activities in the morning. On the other hand, people living with good accessibility of public transportation system have shorter duration on work activity but longer duration for personal business. As for the state-dependent covariates, the beginning time of activity impacts its duration. The results indicate that for home, work, school and social-recreation activities, earlier the beginning time of activities is, longer persists its duration. As for the influence of the duration of previous travel/activity conducted, the results show the positive relation between activity duration and derived travel time. This implies that people spend more time on travel to engage an activity with longer duration.

Second non-travel activity episode (EP5)

The second non-travel activity episode begins near 14:07. The number of significant covariates is less than the first non-travel activity episode. For the effects of socio-demographic covariates, it is not surprisingly to find that workers conduct shorter duration of home activity but it is longer for people of age between 55 to 65 years. For work activity, the result reveals that men have shorter work time in this early afternoon episode, reflecting possibly that different engagement of activities for men occurs usually in the afternoon. On the other hand, it is interesting to find that people of age more than 65 years have longer work time in this episode. For other activities, it shows that workers and couple conduct shorter shopping activity for this episode. Men and couple engage longer personal business activity in the afternoon. For social-recreation activity, it is reasonable to observe that workers have shorter duration since the disavailability of time for this purpose. Inversely, people of age between 25 to 55 years spend longer time on this activity. It is interesting to find that the presence of children has no significant effect for all activities. The effects of transport accessibility indicate that people with car proceed shorter duration for home and work activities. This result implies that the availability of car provides more flexible mobility to conduct another activities and hence reduce the duration in home and work activity at this episode. As for the effects of state-dependent covariates, the result suggests that if the beginning time of home, work and school activities is later in the afternoon, its duration will be shorter. On the other hand, it is not surprising to find the inverse effect is observed for social-recreation activity. The effect of duration of trip previously conducted has similar negative effect for each of activity duration.

Third non-travel activity episode (EP7)
This episode begins around the middle of afternoon (15:00). The effects of covariate are quite different with respect to prior non-travel activity episode. For home activity, it reveals that men have shorter duration. For work activity, workers and people without the presence of children conduct longer duration. As for the school activity, men and workers conduct longer duration. For shopping activities, it is reasonable to find that children conduct less time in shopping. For personal business and social-recreation activities, the result suggests that men conduct longer duration for these two types of activities. The effect of the presence of children reduces the duration of personal business and social-recreation activities. This might imply that the need to accompany children to home and then reduce available time to participate this activity in this episode. As for the urban form covariate, the result suggests that it is not significant for each of activity purposes. As for the effect of transport accessibility, the result indicates that people with car have shorter duration in personal business and social-recreation activities. Similar effect is observed for the proximity of public transportation system on work activities. However, it is worth noting that it has longer duration effect on personal business activity. As for the state-dependent covariate, it suggests that if the beginning time on home, work and school is later, its duration is shorter. The inverse effect is observed for personal business activity. It is interesting to find that the derived travel time is significant for shopping, personal business and socio-recreation activities.

Fourth non-travel activity episode (EP9)
This episode begins around late afternoon (16:45), period of beginning the return home after work. The number of significant covariates is less for each of activity purposes. For socio-demographic covariates, it reveals that workers and young people of age between 15 to 25 years have shorter home activity duration. Couples have longer duration in work and shopping activities. The presence of children has longer activity duration effect on work activity in this episode. It is worth noting that there is no significant effect for the urban form covariate. As for the effect of transport accessibility, people with car have shorter duration for personal business, similar effect observed in the previous non-travel episode. As for the proximity of public transport, the effect is inverse with respect to previous non-travel activity episode. As for the state-dependent covariates, the results suggest that the beginning time has shorter duration effect on activity duration. Note that the estimated parameter for school activity is rather large due to the small sample observed for this transition. Hence the related statistical inference should be proceeded with care. Similar effect of derived travel time on activity duration is also observed for shopping activity.

In summary, it is interesting to note that most cause-specific covariates have variable effects on the duration of activities conducted in different episodes with exceptions of the presence of car and the duration of trip previously conducted.

Effects of covariates for travel episodes

The effects of covariates on travel duration reveal individual's travel time expenditure difference of socio-demographic classes. It also reflects the effects of related covariates on travel time use with respect to different activity purposes. In the following, the effects of covariates on travel time duration are discussed. The results are shown in table 5, drawn from the observations of travel time data conducted in the first 4 travel episodes (EP2, EP4, EP6 and EP8). Note that the medians of the starting time of these episodes are 8:15, 13:35, 14:40 and 16:30. It is expected to find variable episode-specific effects of covariates for each of activity purposes.

First travel episode (EP2)
The first travel episode represents the first trip in the morning for different purpose of activities. The effect of socio-demographic covariates is discussed firstly. The coefficient of gender is

negative for work and social-recreation activities but positive for school activity. It means that men spend longer travel time for work and social-recreation activities, but shorter travel time for school activity. The effect of age on travel time indicates that children of age less than 15 years have shorter travel time for school and shopping activities. This results possibly from the fact that primary/secondary schools are usually located with a shorter distance from residence. For people of age between 25 and 55, the results suggest that they have shorter travel time for shopping and social-recreation activity. As for older people of age between 55 and 65 years, it shows that they have longer travel time for personal business activity. For people of age more than 65 years, the parameter estimate indicates that they have longer travel time for work, a special finding of observed data. Individuals with presence of children spend less travel time for home, school and personal business activities. As for the urban form effects, people living in city center spend less travel time for shopping activities but more travel time for social-recreation activity. People with the availability of car spend less travel time for school and personal business activities. This may imply that the availability of car gives more flexible choice of travel mode and results in the less travel time for these activities. As for the state-dependent covariates, the departure time of the first out-home trip impacts significantly its duration. This finding is reasonable since travel time is largely impacted by the presence of traffic congestion in peak hours of morning. The parameter estimates indicate that for work, school and social-recreation activities, later the departure time is, shorter the travel duration spends on it. However, for personal business activity, the results indicate inverse effect of departure time on the trip duration.

Second travel episode (EP4)
The starting time of the second travel episode reveals that people start their second out-home activities at early afternoon. The parameter estimates of this episode indicate that the effects of socio-demographic covariates impact mainly on school, shopping and personal business. The results indicate that men, people of age than 15 years have shorter travel duration for school activity but it is longer for people of age more than 65 years. For shopping and personal business activities, the parameter estimates indicate that people with the presence of children spend less travel time on these activities. However, women and people of age between 55 and 65 years spend longer travel time for shopping and personal-business activity, respectively. For social-recreation activity, people of age more than 65 years spend shorter travel time for this activity purpose. As for the urban form covariates, the results indicate that people living in city center spend less travel time for shopping activity. This finding may imply that better accessibility in center city for shopping and maintenance activities. As for transport accessibility covariates, the results suggest that people with car ownership have shorter travel time for school activity. On the other hand, public transport proximity shows inverse effect on the travel duration of school activity. Finally, for state-dependent covariates, the parameter estimate is negative for the trip departure time and the duration of previous activity. This result implies that later/longer the departure time/duration of previous activity is, more the travel time is spent on this trip. The possible explanation is that people departing late in the afternoon may join in the second vague of travel demand and result in the increasing travel time. Note that the departure time effect is significant only for home and shopping activities. On the other hand, the effect of previous activity duration is significant for home, personal-business and social-recreation activity.

Third travel episode (EP6)
The third travel episode begins in average at 14:40 (table 4). It reflects the third travel episode in the day. For socio-demographic covariates, it is interesting to find that men spend longer travel time in work and shopping activities. For the effect of age, the results indicate that people of age less than 15 years have shorter travel time on work, school and personal-business activities. Similar effect is found on people of age between 15 to 25 years for work and social-recreation activities. For the other covariates, workers have longer travel duration for shopping and people with the presence of children have shorter travel time in home, shopping and personal-business activities. As for the urban form covariate, the result indicates that people living in city center have longer travel time for social-recreation activity. For transport accessibility covariates, it is

interesting to find that the availability of car is not significant for this trip, resulting possibly from the shorter travel duration of this trip without the interests of using car. On the other hand, individuals with public transportation proximity have longer travel time for work and personal business activities. This result may indicate that workers prefer to use public transportation system for business visit and personal-business in the afternoon. As for the state-dependent covariates, the duration of activity previously conducted has negative effect on travel time spent on trip for next activity. This result is similar as the second trip episode, resulting possibly from the dependency effect of duration process of travel and non-travel activity, constrained combinedly by available time within the context of individual's daily scheduled activity program.

Fourth travel episode (EP8)
This fourth travel episode begins around the return home period after work. The socio-demographic covariates have less significant effects on the duration of this trip. The parameter estimates indicate that women and couple have longer travel time for school, resulting possibly from the additional travel needs for accompanying children to home. For the effects of other covariates, it is reasonable to find that children have less travel time for work and school activity. For other socio-demographic covariates, workers have longer travel time for socio-recreation activity. Couple and people with the presence of children have longer travel time for work. The presence of children has negative effect on the duration of trip for shopping activity. It is interesting to find that there are no significant effects for urban form and transport availability covariates in this evening episode. As for state-dependent covariates, the parameter estimate indicates that the departure time has large positive effect on trip duration for school activity. This result should be explained with caution since there is only 8 observed samples for this transition. For the effect of duration of activity previously conducted, it reveals the negative effect of travel duration, similar to previous travel episode.

To summarize, the parameter estimates reveal that variable effects of covariates on travel time use for each of activity purposes. It also exhibits daily travel time use rhythm with respect to activity type and related covariates. Interestingly, we find that the duration of trip depends on its trip purpose, starting time-of-day and also the duration of activity conducted in previous episode. The result coincides with previous empirical study (*14*).

Table 3 Model fit statistics

Episode	Current state	Next state	Log-liklihood	Test	-2Log-Likelihood Ratio	DF	P value
EP 1	Home	Trip 1	--			--	--
EP 2	Trip 1	Activity 1	-15080.26			18	
			-15034.33	2 vs 1	91,858	22	<0.0001
EP 3	1 Home	Trip 2	-100.50			3	
			-97.85	2 vs 1	5,297	4	0.0122
			-97.14	3 vs 2	1,422	4	--
	2 Work	Trip 2	-5382.18			3	
			-5350.48	2 vs 1	63,405	5	<0.0001
			-5341.73	3 vs 2	17,497	7	<0.0001
	3 School	Trip 2	-3484.03			5	
			-3472.75	2 vs 1	22,554	7	<0.0001
			-3463.67	3 vs 2	18,178	6	<0.0001
	4 Shopping	Trip 2	-2480.89			3	
			-2479.95	2 vs 1	1,876	4	0.1140
			-2455.17	3 vs 2	49,564	4	--
	5 Personal business	Trip 2	-2213.90			7	
			-2212.70	2 vs 1	2,397	8	0.0770
			-2204.83	3 vs 2	15,745	8	--

Table 3 Model fit statistics (Continuous)

Episode	Current state	Next state	Log-liklihood	Test	-2Log-Likelihood Ratio	DF	P value
	6 Social-Recreation	Trip 2	-1422.67			5	
			-1417.62	2 vs 1	10,102	6	0.0008
			-1411.37	3 vs 2	12,492	5	0.0002
EP 4	Trip 2	Activity 2	-15239.82			15	
			-15192.53	2 vs 1	94,564	17	<0.0001
			-15169.01	3 vs 2	47,051	17	--
EP 5	1 Home	Trip 3	-10889.39			6	
			-10751.94	2 vs 1	274,91	4	<0.0001
			-10748.34	3 vs 2	7,189	5	0.0040
	2 Work	Trip 3	-1183.98			3	
			-1162.34	2 vs 1	43,268	4	<0.0001
			-1160.89	3 vs 2	2,899	5	0.0553
	3 School	Trip 3	-61.15			2	
			-60.15	2 vs 1	1,987	2	--
			-56.20	3 vs 2	7,901	3	0.0027
	4 Shopping	Trip 3	-1082.03			4	
			-1082.03	2 vs 1	0	4	--
			-1067.77	3 vs 2	28,527	3	<0.0001
	5 Personal business	Trip 3	-1039.93			4	
			-1036.44	2 vs 1	6,979	5	0.0046
			-1034.17	3 vs 2	4,535	5	--
	6 Social-Recreation	Trip 3	-1007.28			4	
			-1007.28	2 vs 1	0	4	--
			-1002.71	3 vs 2	9,148	3	0.0013
EP 6	Trip 3	Activity 3	-6971.91			12	
			-6941.16	2 vs 1	61,495	18	<0.0001
			-6905.08	3 vs 2	72,152	23	<0.0001
EP 7	1 Home	Trip 4	-2146.49			6	
			-2089.19	2 vs 1	114,588	2	<0.0001
			-2089.19	3 vs 2	0	2	--
	2 Work	Trip 4	-1063.44			2	
			-1054.70	2 vs 1	17,47	4	<0.0001
			-1054.70	3 vs 2	0	4	--
	3 School	Trip 4	-367.99			3	
			-353.27	2 vs 1	29,447	3	--
			-353.27	3 vs 2	0	3	--
	4 Shopping	Trip 4	-811.75			2	
			-811.75	2 vs 1	0	2	--
			-805.88	3 vs 2	11,745	2	--
	5 Personal business	Trip 4	-1136.92			4	
			-1135.52	2 vs 1	2,814	5	0.0582
			-1132.66	3 vs 2	5,72	7	0.0286
	6 Social-Recreation	Trip 4	-1421.31			4	
			-1421.31	2 vs 1	0	4	--
			-1414.02	3 vs 2	14,573	5	<0.0001
EP 8	Trip 4	Activity 4	-5779.14			11	
			-5776.39	2 vs 1	5,494	10	0.0109
			-5736.30	3 vs 2	80,184	13	<0.0001

Table 3 Model fit statistics (Continuous)

Episode	Current state	Next state	Log-liklihood	Test	-2Log-Likelihood Ratio	DF	P value
EP 9	1 Home	Trip 5	-4043.89			4	
			-3948.23	2 vs 1	191,322	3	<0.0001
			-3948.23	3 vs 2	0	3	--
	2 Work	Trip 5	-392.42			2	
			-390.08	2 vs 1	4,68	3	0.0177
			-389.01	3 vs 2	2,137	4	0.0937
	3 School	Trip 5	-9.52			1	
			-5.18	2 vs 1	8,678	2	0.0017
			-5.18	3 vs 2	0	2	--
	4 Shopping	Trip 5	-340.77			3	
			-340.77	2 vs 1	0	3	--
			-334.53	3 vs 2	12,468	3	--
	5 Personal business	Trip 5	-509.61			1	
			-509.61	2 vs 1	0	1	--
			-509.61	3 vs 2	0	1	--
	6 Social-Recreation	Trip 5	-467.39			3	
			-467.39	2 vs 1	0	3	--
			-467.39	3 vs 2	0	3	--

Table 4 Summary statistics of beginning time of episodes

Episode	Median	Std. Dev.
EP2 (travel)	495 (08hr15min)	165.85
EP3 (non-travel activity)	510 (08hr30min)	166.37
EP4 (travel)	815 (13hr35min)	201.88
EP5 (non-travel activity)	847 (14hr07min)	206.56
EP6 (travel)	880 (14hr40min)	185.17
EP7 (non-travel activity)	900 (15hr00min)	188.47
EP8 (travel)	990 (16hr30min)	183.26
EP9 (non-travel activity)	1005(16hr45min)	184.87

Remark: in minute

CONCLUSIONS

This study applies multi-state semi-Markov model for daily travel-activity duration estimation. The results of estimates suggest that the duration of activity depends not only on its type, beginning time-of-day and the duration of travel/activity previously conducted. It reveals also the temporal rhythm of the travel and activity duration patterns during different periods of a day. The empirical estimates of the baseline hazard are different for each of activity purposes across episodes. The cross-episode comparison of these baseline hazard profiles shows interesting variation of temporal rhythm of activity duration, which confirms again the strong dependency effect of the starting time-of-day on the travel-activity duration pattern formation.

The model-fit test suggests that the incorporation of state-related covariates into hazard function is important to estimate the duration of activity. Based on the log-likelihood value, the final model incorporating aforementioned state-dependent covariates performs best over the other models. For transition hazard estimation, the Cox regression model is applied for each of one-step transitions in individual's travel-activity chain. We test the proportionality assumption of the Cox regression model based on the cumulative sums of residuals for all related covariates in the transitions across episodes. It is interesting to find that the proportionality assumption is valid for almost each of covariates included in the final model except only the covariate of “the logarithm of entering time of activity” in home-trip and that in school-trip transition in episode 7 and 9, respectively.

The results of parameter estimations indicate that people spend more time on travel if the duration of activity to participate is longer. Moreover the covariates in terms of socio-demographic characteristics, urban form and transport accessibility perform different effects on activity duration during different episode of a travel-activity chain. The empirical estimates of baseline hazard suggests that the home activity duration after the first trip indicates a duration pattern of short sojourn time of 25 minutes and a longer one of 200 minutes in the morning. For work activity, the results indicate that a general rupture pattern of work duration with a rhythm of 10 to 15 minutes during the later afternoon period. For school activity, it is interesting to find a pattern of spikes in the baseline hazard at 20, 70-80, 130 and 200 minutes, representing interesting duration cadence of school activities. For shopping activity, the result indicates the later begins the shopping activity, shorter is its duration tends to be. For personal business activity, the result shows a general rupture rhythm of 10-15minutes during the first 2-hour period in the afternoon. On the other hand, similar rupture rhythm of 10-15minutes is observed during a wider range of 4 hours for social-recreation activity.

The proposed multi-state semi-Markov model allows one to investigate the dependency effect of travels/activities conducted in a travel-activity chain. It provides also useful insight on the effect of covariates on travel/activity duration. Moreover, the proposed model can be used to predict individual’s travel-activity patterns with respect to related covariate settings. It can also be used for generating activity program of homogeneous socio-demographic classes for activity-based travel demand analysis. The future extensions of current study may incorporate the effect of heterogeneity in hazard function specification and compared the estimate results over different socio-demographic groups.

REFERENCES

1. Chapin, F.S. The Study of Urban Activity Systems. *Urban Land Use Planning*. University of Illinois Press. Urbana. 1965.
2. Hemmens, G.C. Analysis and simulation of urban activity patterns. *Socio-Economic Planning Sciences*. Vol. 4. No. 1. 1970. pp. 53-66.
3. Arentze, T. and H.J.P. Timmermans. A learning-based transportation oriented simulation system. *Transportation Research Part B*. Vol. 38. No. 7. 2004. pp. 613-633.
4. Bhat, C.R., S. Srinivasan and K.W. Axhausen. An analysis of multiple interepisode durations using a unifying multivariate hazard model. *Transportation Research Part B* 39. 2005. pp.797-823.
5. Lee, Y., M. Hickman and S. Washington. Household type and structure. time-use pattern. and trip-chaining behavior. *Transportation Research Part A*. Vol. 41. 2007. pp.1004–1020.
6. Recker. W.W., M.G. McNally and G.S. Root. A model of complex travel behavior: Part 1: Theoretical development. *Transportation Research A*. Vol. 20. 1986. pp.307–318.
7. Recker, W.W., M.G. McNally and G.S. Root. A model of complex travel behavior: Part 2: An operational model. Transportation Research A. Vol. 20. 1986. pp.319–330.
8. Ettema, D., A. Borgers and H.J.P. Timmermans. Competing risk hazard model of activity choice. timing. sequencing. and duration. In: Transportation Research Record: Journal of the Transportation Research Board. No. 1493. Transportation Research Board of the National Academies. Washington. D.C. 1995. pp.101–109.
9. Bhat, C.R. A hazard-based duration model of shopping activity with nonparametric baseline specification and nonparametric control for unobserved heterogeneity. *Transportation Research Part B*. Vol. 30. No. 3. 1996. pp.189–207.

10. Lee, B. and H.J.P. Timmermans. A latent class accelerated hazard model of activity episode durations. Transportation Research Part B. Vol. 41. No. 4. 2007. pp.426-447.
11. Bhat, C.R. A generalized multiple durations proportional hazard model with an application to activity behavior during the evening work-to-home commute. Transportation Research Part B. Vol. 30. No. 6. 1996. pp. 465–480.
12. Srinivasan, K.K. and Z. Guo. Analysis of trip and stop duration for shopping activities: Simultaneous hazard duration model system. In *Transportation Research Record: Journal of the Transportation Research Board. No* 1854. Transportation Research Board of the National Academies. Washington. D.C. Paper No. 03-4499. 2007.
13. Pendyala, R. M. and C.R. Bhat. An exploration of the relationship between timing and duration of maintenance activities. *Transportation*. Vol. 31. 2004. pp. 429–456.
14. Popkowski Leszczyc. P.T.L. and H. Timmermans. Unconditional and conditional competing risk models of activity duration and activity sequencing decisions: An empirical comparison. *Journal of Geographical Systems*. Vol. 4. 2002. pp. 157–170.
15. Lee. B., H.J.P. Timmermans. Z. Junyi and A. Fujiwara. A generalized multi-state hidden markov extension model for multi-state activity episode duration. *Proceeding of the 11th international conference on travel behavior research*. Kyoto. 2006.
16. Hougaard P. Multi-state Models: A Review. *Lifetime Data Analysis*. Vol. 5. No. 3. 1999. pp. 239-264.
17. Dabrowska, D.M., G.W. Sun and M.M. Horowitz. Cox regression in a Markov renewal model: an application to the analysis of bone marrow transplant data. *Journal of the American Statistical Association*. Vol. 89. No. 427. 1994. pp. 867–877.
18. Putter, H., J. van der Hage, G. H. de Bock, R. Elgalta and C.J.H. van de Velde. Estimation and prediction in a multi-state model for breast cancer. *Biometrical journal*. Vol. 48. 2006. pp. 366-380.
19. Cox, D.R. Regression Models and Life Tables. *Journal of the Royal Statistical Society B*. Vol. 34. 1972. pp. 187-200.
20. Kalbfleisch, J.D. and R.L. Prentice. *The Statistical Analysis of Failure Time Data*. New York: John Wiley & Sons. Inc. 2nd Edition, 2002.
21. Breslow, N.E. Covariance Analysis of Censored Survival Data. *Biometrics*. Vol. 30. 1974. pp. 89-99.
22. Efron. B. The Efficiency of Cox's Likelihood Function for Censored Data. *Journal of the American Statistical Association*. Vol. 72. 1977. pp. 557-565.
23. Therneau, T.M., P.M. Grambsch and T.R. Fleming. Martingale-based residuals for survival models. *Biometrika* Vol. 77. 1990. pp. 147-160.
24. Spiekerman, C.F. and D.Y. Lin. Checking the marginal Cox model for correlated failure time data. *Biometrika*. Vol. 83. 1996. pp. 143-156.
25. Lin, D.Y., L.J. Wei and Z. Ying. Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. *Biometrika*. Vol. 80. 1993. pp. 557-572.
26. Lin, D.Y., L.J. Wei and Z. Ying. Model-Checking Techniques Based on Cumulative Residuals. *Biometrics.* Vol. 58. 2002. pp. 1-12.
27. Dabrowska, D.M. and W.T. Ho. Estimation in a semiparametric modulated renewal process. *Statistica Sinica* Vol. 16. 2006, pp. 93-119.

Table 5 Covariates parameter estimates and standard derivation

Episode	State transition	Gender	Age15	Age25_55	Age55_65	Age65	Prof	H_type	H_location	N_children	N_Car	PT_proximity	log(t_i)	Log (duration of travel/activity conducted in previous episode)
EP 1	H-Tr1													
EP 2	Tr1-H									0.24[b](0.11)				
	Tr1-W	-0.13[a](0.06)				-0.69[a](0.35)						-0.06[a](0.03)	0.67[a](0.15)	
	Tr1-Sc	0.16[a](0.08)	0.69[a](0.09)							0.05[a](0.03)	0.17[a](0.05)	-0.06[b](0.03)	2.64[a](0.51)	
	Tr1-Sh		0.63[a](0.18)	0.31[a](0.08)					0.27[a](0.08)					
	Tr1-PB				-0.23[a](0.11)					0.12[a](0.04)	0.15[a](0.06)		-0.38[a](0.15)	
	Tr1-SR	-0.18[b](0.11)		0.24[a](0.12)					-0.26[a](0.12)				0.55[a](0.16)	
EP 3	H-Tr2					2.26[a](0.95)			0.73[a](0.37)				1.38[a](0.47)	-0.33[b](0.18)
	W-Tr2	-0.22[a](0.07)							0. 26[a](0.07)	0. 06[b](0.03)	0.09[a](0.04)	0.05(0.03)	1. 46[a](0.17)	-0.16[a](0.04)
	Sc-Tr2	-0.12(0.08)				5.03[a](1.07)	0.61[a](0.31)	-0.54[a](0.13)					2. 18[a](0.45)	-0. 23[a](0.05)
	Sh-Tr2			0.18[b](0.10)				-0.24[a](0.12)		-0.11[a](0.05)				-0.41[a](0.06)
	PB-Tr2		0.75[a](0.32)	0.32[a](0.14)			0.20[b](0.12)		0.19[b](0.11)	0.07(0.05)	0.17[a](0.06)	-0.09[a](0.05)		-0.23[a](0.06)
	SR-Tr2					-0.26(0.13)		-0.35[a](0.14)	0.32[a](0.13)				0. 63[a](0.20)	-0. 23[a](0.05)
EP 4	Tr2-H												-0.48[a](0.15)	-0.21[a](0.03)
	Tr2-W													
	Tr2-Sc	0.13[b](0.08)	0.67[a](0.08)		-1.80[a](0.30)						0.10[a](0.05)	-0.09[a](0.03)		
	Tr2-Sh	-0.20[a](0.08)							0.25[a](0.07)	0.11[a](0.03)			-1.04[a](0.28)	
	Tr2-PB				-0.33[a](0.13)					0.06(0.04)				-0.08[a](0.04)
	Tr2-SR					0.32[a](0.12)							0.58(0.39)	-0.18[a](0.06)
EP 5	H-Tr3				-0.27[a](0.08)		0.23[a](0.05)				0.10[a](0.03)		2.6[a](0.17)	-0.08[a](0.03)
	W-Tr3	0.30[a](0.13)				-0.78(0.52)					0.14[b](0.08)		1.84[a](0.25)	-0.13[b](0.07)
	Sc-Tr3			-1.30[b](0.69)									3.14[a](0.89)	-0.74[a](0.24)
	Sh-Tr3						0.26[a](0.13)	0.40[a](0.18)						-0.45[a](0.08)
	PB-Tr3	-0.32[a](0.14)						-0.32[b](0.17)				-0.08[b](0.05)	-0.74[a](0.29)	-0.17[a](0.07)
	SR-Tr3			-0.75[a](0.20)			0.87[a](0.20)							-0.25[a](0.07)

Remark: t_i is the entering time (in minute) of current state i

( ) : Standard deviation

a : Significant at the 0.05 level. b : Significant at the 0.1 level.

H:home, W: work, Sc: school, Sh: shopping, PB: Personal business. SR: Social-Recreation

Table 5 Covariates parameter estimates and standard derivation (continue)

Episode	State transition	Gender	Age15	Age15_25	Age25_55	Age55_65	Prof	H_type	H_location	N_children	N_Car	PT_proximity	log(t_i)	log(duration of travel/activity conducted in previous episode)
EP 6	Tr3-H									0.32^{a}(0.15)			-0.96^{a}(0.25)	-0.25^{a}(0.04)
	Tr3-W	-0.22^{a}(0.09)	3.51^{a}(0.50)	0.34^{b}(0.20)			0.19(0.11)					-0.09^{a}(0.03)		-0.24^{a}(0.05)
	Tr3-Sc		0.26^{a}(0.10)					-0.43^{b}(0.20)						
	Tr3-Sh	-0.23^{a}(0.09)					-0.22^{a}(0.11)			0.23^{a}(0.05)				-0.13^{b}(0.05)
	Tr3-PB		0.55^{a}(0.25)							0.10^{a}(0.04)		-0.07^{a}(0.03)		-0.18^{a}(0.04)
	Tr3-SR			0.53^{a}(0.20)					-0.46^{a}(0.11)				1.02^{a}(0.37)	-0.09^{b}(0.05)
EP 7	H-Tr4	0.15(0.10)											3.92^{a}(0.38)	
	W-Tr4						-1.36^{a}(0.46)			0.10^{b}(0.06)		0.12^{b}(0.07)	1.48^{a}(0.36)	
	Sc-Tr4	-0.35^{b}(0.21)					-1.89^{a}(0.70)						5.51^{a}(0.95)	
	Sh-Tr4		0.79^{a}(0.31)											-0.35^{a}(0.09)
	PB-Tr4	-0.29^{a}(0.13)			0.21(0.14)					0.18^{a}(0.06)	0.30^{a}(0.08)	-0.17^{a}(0.06)	-0.45(0.30)	-0.15^{a}(0.07)
	SR-Tr4	-0.27^{a}(0.12)				-0.32(0.21)				0.11^{a}(0.05)	0.18^{a}(0.08)			-0.23^{a}(0.06)
EP 8	Tr4-H													-0.26^{a}(0.03)
	Tr4-W		3.30^{a}(0.46)					-0.25^{a}(0.11)		-0.08^{a}(0.04)				
	Tr4-Sc	0.32^{a}(0.13)	0.21^{a}(0.10)					-0.45^{a}(0.19)					13.28^{a}(4.23)	-1.81^{a}(0.40)
	Tr4-Sh									0.14^{a}(0.05)				
	Tr4-PB													-0.22^{a}(0.05)
	Tr4-SR						-0.48^{a}(0.17)							-0.33^{a}(0.09)
EP 9	H-Tr5			0.37^{a}(0.11)			0.28^{a}(0.08)						5.07^{a}(0.40)	
	W-Tr5							-0.63^{a}(0.29)		-0.25^{a}(0.09)			1.08^{b}(0.58)	0.19(0.13)
	Sc-Tr5											-1.87(1.26)	47.72^{b}(28.6)	
	Sh-Tr5			0.70^{a}(0.35)				-0.49^{b}(0.27)						-0.47^{a}(0.12)
	PB-Tr5										0.21^{b}(0.13)			
	SR-Tr5	-0.28(0.19)				-0.49^{b}(0.27)						0.17^{a}(0.07)		

Remark: t_i is the entering time (in minute) of current state *i*

( ) : Standard deviation

a : Significant at the 0.05 level. b : Significant at the 0.1 level.

H:home, W: work, Sc: school, Sh: shopping, PB: Personal business. SR: Social-Recreation