Comparison of Clustered-based MNL and Duration Models in Departure Time Choice Modelling

Shahriar Afandizadeh ZARGARI*, Farshid SAFARI

Abstract: One of the important decisions for travellers about each trip is the time that they depart from their point of origin. The aggregation of the departure times of all travellers forms the pattern of a temporal distribution of trips in a day and also peak periods of a day. In this paper, we applied the multinomial logit (MNL) on a choice set which was derived from cluster analysis, and we also made use of the duration models to estimate departure time for home-based work trips. For the duration models, the Kaplan-Meier and Cox proportional methods are used and then, the results are compared to cluster-based MNL model. This study is conducted in Mashhad city which has a population of about 2.7 million people. The results show the traveller's job and traveller's selected mode have a significant effect on his departure time choice. Comparison of the predictability power of these two modelling approaches indicates that the cluster-based MNL model in this case study is the preferable model.

Keywords: clustered-based MNL; cox proportional method; Departure Time Choice (DTC); duration models; K-mean clustering

1 INTRODUCTION

One of the important traveller decisions about each trip is the time that a traveller departs from his/her origin. Each person chooses the departure time independently and this choice is originated in experienced travel time, travel costs, marital status, family obligations, income, occupation and the flexibility of work, travel mode and the importance of activity [1]. The aggregation of the departure times of all travellers forms the pattern of a temporal distribution of trips in a day and also peak periods of a day. Particularly, the commuters' choice of departure time from home, which is a crucial factor in determining how congestion will change through the morning peak period [2].

In almost all urban areas, the temporal trip pattern shows that the main portion of demand start their trips in a short period and in this duration, public and private network capacity will be inadequate and therefore, congestion and delays will be inevitable. Additionally, analysis of various scenarios at different times in a day is essential for economic and environmental evaluation of decisions in the transportation planning. The preferred arrival times are commonly concentrated in a short period, which forms peak hours. Demand management policies try to spread the peak over a longer period and make a trade-off between arriving early/late at the destination and thus spending less time in congestion or arriving on time but spending more time in congestion [3]. For performing this analysis, estimation of demand in different time intervals is considered.

Thus, modelling departure time choice is an important part of the travel demand modelling process. The major applications of this model are analysis and evaluation of different demand management policies (modification in the trip pattern) [4-6], dynamic assignment models, the response of travellers to ATIS systems, air quality modelling and activity-based models [7].

This paper is organized as follows: In section 2, we review the literature on departure time choice, then section 3 provides an overview of the duration model and cluster-based MNL model used in the analysis and in section 4 we present empirical results. Finally, section 5 emphasizes the findings of this study.

2 LITERATURE REVIEW

Different models are used for selecting the departure time including multinomial logit model (MNL); nested logit model; cross-nested logit model; mixed logit model; continuous time model; and ordered generalized extreme value (OGEV) model. Two key challenges are involved in DTC modelling. Small and Cascetta et al. [8, 9] employed multinomial logit models to select a departure time. In 1987, Small [10] proposed an ordered generalized extreme value model, in which several departure time periods were considered in a nested design. Hendrickson et al. [11] used the data collected from Pittsburgh to examine flexibility of departure time choices for work trips. In these applications of the MNL model, a day is divided into a number of periods, and the MNL model is used to select a choice from the set of choices (periods). In the MNL model, systematic utility functions are defined as a function of socioeconomic attributes and variables related to trip purpose. Although the MNL model overlooks the similarities or correlations among adjacent periods and violates the IID assumption, it is still popular due to its closed form and extensive use [12].

Bhat [13] referred to the importance of the mode and departure time choices although the priority of researchers modelling travel demand was mode choice and little attention has been paid to the DTC. He introduced a nested structure by considering the mode choice in the higher level and the DTC in the lower level. The MNL model was used for mode choice, while the ordered generalized extreme value (OGEV) model was used to consider ranks of choices and selection of departure time. In another study, Bhat & Steed [4] used the "hazard-based duration model" to develop a continuous model of DTC for urban shopping trips. Habib et al. [7] used the "hazard-based duration model" to select the departure time. Application of this model for DTC does not necessarily represent behaviours.

Ettema et al. [14] developed a DTC model based on activity plans. They used the utility of participation in activities to model the DTC. Lim et al. [15] investigated the logit-based combined departure time and dynamic stochastic user equilibrium assignment (DDSUE) problem. Ben-Elia et al. [16] proposed a DTC model based on the

latent preferential arrival time notion. Using this model, they developed a modelling framework for the calculation of the DTC by assuming a latent class of the preferential arrival time. Sasic & Habib [17] proposed a discrete choice model with a latent choice set. This model is one type of generalized extreme value (GEV) model.

To conclude, various modelling approach has been used in DTC models which are shown in Fig. 1 that shows the two major modelling approaches are discrete choice models and continuous models. Furthermore, some researchers have applied joint models, activity-scheduling models and changing departure time models for different problem definitions. The two major modelling approaches have some strengths and weaknesses. Regarding random utility maximization (based on rational decision maker), discrete choice modelling could reflect traveller's behaviour and therefore these models will predict the effect of change of major variables in any time horizon. On the contrary as a weakness, based on continuous nature of time, choice set generation will be a challenge. The time discretization leads to some intervals that form the choice set and travellers choose one of them but the departure times near the edge of two adjacent intervals are highly correlated and do not match with independence from irrelevant alternatives (IIA) assumption of discrete choice models.
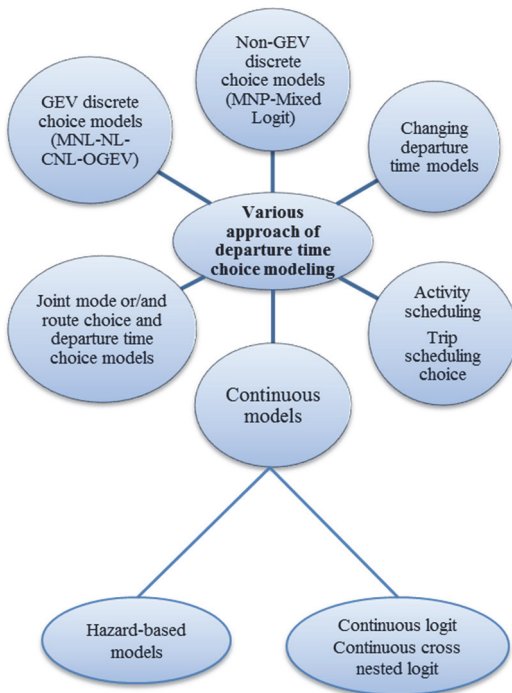

Figure 1 Various modelling approach in departure time choice models

3 METHODS
3.1 Cluster-based Multinomial Logit Modelling

In the discrete choice models, each choice is favourable to a certain degree for a person and the choice with the highest level of utility will be selected. The utility function of the person named n that corresponds to choice i in the C_n choice set is determined through the following equation.

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

where, V_{in} is the deterministic term of the utility and ε_{in} is the random term of the utility, which denotes the uncertainty caused by the limited power of the analyst. The deterministic term of the utility is determined in modelling and the effect of different qualities on the selection of choices is identified.

In other research, we evaluated various clustering methods for choice set generation in MNL modelling of departure time [18]. The assessed clustering methods in that study are K-means, K-medoids, EM clustering, and Hierarchical Ward clustering method. The result of that study showed estimated models based on K-means clustering method have more accurate prediction, so K-means clustering method is applied in this research.

3.2 Duration Models

The duration models (sometimes called "hazard-based" models or "survival" models) have three main characteristics related to data. First, the dependent variable is the waiting time until the occurrence of a well-defined event. Second, observations are censored, in the sense that for some units the event of interest has not occurred at the time the data are analysed; and third, there are explanatory variables whose effect on the waiting time we wish to assess or control. These characteristics match departure time nature when we assume departing from the origin as an event. In this approach, departure time (T) is defined as a continuous random variable with probability density function $f(.)$ and cumulative density function $F(t)$ giving the probability that departure event has occurred by duration t. It will often be convenient to work with the complement of the c.d.f, the survival function.

$$S(t) = \Pr\{T \ge t\} = 1 - F(t) = \int_t^\infty f(x)\,\mathrm{d}x \tag{2}$$

An alternative characterization of the distribution of T is given by the hazard function, or instantaneous rate of occurrence of the event, defined as

$$\lambda(t) = \lim_{\mathrm{d}t \to 0} \frac{\Pr\{t \le T < t + \mathrm{d}t \,|\, T \rangle t\}}{\mathrm{d}t} \tag{3}$$

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\mathrm{d}}{\mathrm{d}t} \log S(t) \tag{4}$$

A parametric model based on the exponential distribution may be presented as follows:

$$\lambda_i(t) = \exp(\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}) \tag{5}$$

In this case, the constant α represents the log-baseline hazard since $\log \lambda_i(t)$ when all the x's are zero. A large family of models introduced by Cox (1972) focuses directly on the hazard function. The Cox proportional hazards model is a semi-parametric model where the baseline hazard $\alpha(t)$ is allowed to vary with time:

$$\log \lambda_i(t) = \alpha(t) + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} \tag{6}$$

$$\lambda_i(t) = \lambda_0(t) \cdot \exp\left(\beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}\right) \qquad (7)$$

If all of the x's are zero the second part of the above equation equals 1, $\lambda_i(t) = \lambda_0(t)$ so the term $\lambda_0(t)$ is called the baseline hazard function. With the Cox proportional hazards model, the outcome is described in terms of the hazard ratio [19]. In this study, we applied Kaplan-Meier (a nonparametric method) and Cox proportional method (a semi-parametric) for modelling and analysis of departure time.

4 RESULTS
4.1 Data Preparation and Explanatory Analysis

The data used for this analysis is the 2008 survey conducted in Mashhad City with a population of about 2.7 million. This city is the second most populous metropolitan city in Iran [20, 21]. The sample size was about 3% of the population and household, household members, vehicle, trips and activities data were gathered. About 62 percent of Mashhad's residents are students and workers which comprise 51 percent of total daily trips and 84 percent of morning peak period trips. Therefore this group has a major role in forming morning peak periods. In this research, we considered those trips whose goals were job, education, medical care and going to offices. In addition, this study assumes parents choose departure time for children under 10 years old and therefore we filter trips of under 10-year-old travellers [22]. The departure time of travellers that was used for this analysis was between midnight and 13:00. The final sample contains 12467 individuals, 70 percent of which was used for model calibration and the rest was used for forecasting performance evaluation. Appendix1 shows descriptive information on work departure time which is based on the method of Kaplan-Meier.

4.2 Modelling process and Empirical Results
4.2.1 Cluster-based MNL Model

In this study, we used the result of other research from authors in which various clustering methods were evaluated for choice set generation. The result of that study showed that estimated models based on K-means clustering method have more accurate prediction. In the next step, we use different combinations of variables in modelling and compare their predictability power. Finally, the best model is presented in Tab. 1. In MNL modelling, variables, namely gender, age, profession, vehicle type, vehicle ownership, education level, and travel time are significant. Values of the t-test are shown for each of the estimated values in the table; moreover, the value of ρ_0^2 is shown at the end of the table.

Table1 Cluster-based MNL model for K-means clustering method

	Predefined intervals					
	00:00-6:10	6:10-7:15	7:15-8:18	8:18-9:30	9:30-10:45	10:45-13:00
FAMILY M					−0.15971	
					(−2.80)	
GENDER	0.37355					
	(2.90)					
JOB G1		−1.80974		−0.53302		−0.53302
		(−10.51)		(−5.23)		(−5.23)
JOB_G2	−0.78562	−1.80426		−0.24251		
	(−5.22)	(−9.70)		(−2.10)		
JOB_G3	−0.58602	−2.89328				−0.58602
	(−4.02)	(−14.96)				(−4.02)
EDUCATED		0.28489				
		(3.89)				
CAR OWN	−0.33386					
	(−3.61)					
SC MODE				−1.66956		
				(−2.26)		
PR MODE	−0.34566				−0.34566	
	(−3.80)				(−3.80)	
PU MODE		0.17971				
		(2.66)				
AC MODE	−0.56853					0.54963
	(−3.98)					(2.19)
AGE18P		0.35969				2.14911
		(2.12)				(8.43)
TRAVEL T	1.55774				0.62063	
	(12.58)				(2.97)	
Constant		3.17750	0.86423	1.09339	0.07704	0.35886
		(14.04)	(5.22)	(6.20)	(0.25)	(1.60)
Log Likelihood at Convergence ($N = 8727$) = −6166.78						
Log Likelihood at Constant ($N = 8727$) = −7259.38						
Log Likelihood at Zero ($N = 8727$) = −8975.74						
$\bar{\rho}_0^2$	0.310					
AIC	1.419					
Prediction	($N = 3740$) = 58.63%					

As we expected results of coefficient (parameter) estimation show traveller's job has a significant effect on departure time choice. For example, salespersons show a tendency to depart between 7:15 and 10:45, also professors, physicians, and teachers are prone to start their trips after 7:15. This is predictable because of different start

time and activity duration for various jobs. Moreover, the effect of selected mode on departure time choice is unsurprising. The public transport users are more likely to depart in a period between 7:15 and 8:18. The pedestrians and bicycle riders are more reluctant to depart before 6:10 than the period between 6:10 and 10:45 and they are more inclined to exit homes between 10:45 and 13:00. The long travel times will dictate choosing time periods before 6:10 for morning work shifts and the period of 9:30 to 10:45 for afternoon shifts that are shown with positive coefficient in Tab. 1. In this research, the multinomial logit models were estimated in Nlogit 5.0 (Econometric Software, Inc. NY, USA.) [23]

4.2.2 Semi Parametric Model: Cox Proportional Hazards Model

In this section, we introduce Cox proportional modelling steps (as a semiparametric survival model) briefly. These steps are the selection of covariates, fitting a multivariate model, considering interaction variables and finally checking Cox proportional model assumptions.

For the selection of major covariates, a thorough univariate analysis of the association between departure time and all important covariates has been conducted. The result of this analysis revealed that the gender and use of private car variables were not effective in departure time but other variables have a significant effect. The major covariates in departure time are the use of public mode, walking and riding bicycle, travel time, the age of traveller, being a parent, and four groups of traveller's jobs.

Next, various multivariate models are fitted and the best one was selected. This model contains walking or riding bicycle dummy variable, travel time variable, the age of travellers, being a parent dummy variable, and two groups of travellers' jobs. In the next step, the effect of various interactive variables on the model is investigated and multiplicative variable of travel time to two job groups is entered into the model. The concluding model is presented in Tab. 2.

Table 2 The estimated covariate effects

Variable	Coefficient estimate	Standard error	z	$Pr(>|z|)$
AC_MODE	−0.24740	0.02927	−8.453	$< 2e^{-16}$
PR_MODE	−0.14706	0.02884	−5.099	$3.41e^{-07}$
TRAVEL_T	0.33608	0.04419	7.605	$2.85e^{-14}$
JOB_G2	−0.18547	0.05513	−3.364	0.000768
JOB_G3	−0.41003	0.05944	−6.898	$5.26e^{-12}$
Parent	0.24883	0.02618	9.505	$< 2e^{-16}$
Interaction-1	−0.28144	0.10787	−2.609	0.009079
Interaction-2	−0.18858	0.08208	−2.297	0.021592
Log likelihood function	−70950.08			
Restricted log likelihood	−71231.93			
Chi squared (8 d.f.)	563.7			
Significance level	.00000			
Hosmer and Lemeshow R-squared	0.063			
Wald test (8 d.f.)	575. 8	$p = 0$		
Score (logrank) test (8 d.f.)	585.6	$p = 0$		

Once a suitable set of covariates has been identified, we checked each covariate to ensure that the proportional hazards assumption is valid. As in all regression analyses, some sort of measure analogous to R^2 is needed. Schemper and Stare show that there is not a single and simple measure to assess the goodness-of-fit of a proportional hazards regression model [24]. Hosmer & Lemeshow recommend the following as a summary statistic for goodness of fit [19]:

$$R_M^2 = 1 - \exp\left[\frac{2}{n}\left(L_0 - L_M\right)\right] \qquad (8)$$

where L_0 is the log partial likelihood for the intercept-only model, L_M is the log partial likelihood for the fitted model, and n is the number of cases included in our analysis.

The positive value of the coefficient for "parent" variable shows that parents depart earlier in order to accompany their children to school or day-care. Also, the positive value of "travel time" coefficient shows longer travel time increase likelihood of sooner departure time due to the importance of on-time arrival for work trips. The application of Cox method on 30% of data approximately matched with the application of Cox method on 70% of data. The model predictions were different in comparison with observed data, but the application of Kaplan Meier method on data represents a more accurate prediction. All variables used in MNL and Duration models are described in Tab. 3.

Table 3 Description of variables used in MNL and Duration models

Variable	Description
FAMILY_M	The number of the family members.
GENDER	Male equal one and female zero.
EDUCATED	Anyone who finishes high school equals one and others zero.
CAR_OWN	Families with one or more than one car equal one and others zero.
SC_MODE	Families that use school transport service equal one and others zero.
PR_MODE	Trips that use a private car or motorcycle equal one and others zero.
PU_MODE	Trips that use a public transport equal one and others zero.
AC_MODE	Travelers who use a bicycle or walking in their trip equal one and others zero.
AGE18P	Traveler older than 18 equal one and otherwise zero.
JOB_G1	Travelers who are a workman or army equal one and others zero.
JOB_G2	Travelers who are a teacher, professor, physician or nurse equal one and others zero.
JOB_G3	Travelers who are salespersons equal one and others zero.
Parent	Travelers who are parent equal one and others zero.
TRAVEL_T	The travel time of the trip.

4.2.3 Comparison of Cluster-based MNL Model and Duration Model

In order to find the appropriate model, we compared predictability of Cluster-based MNL and duration model. For this purpose, we used 30% of data which have not been used in model calibration. The Cluster-based MNL model predicts a probability of departure time in each one of 6 intervals, however, duration model forecasts departure time continuously. In order to provide the possibility of comparing two models, we integrated duration model results in 6 predefined intervals in the cluster-based MNL model. The probability of each interval is predicted for each person based on estimated model and then using Monte Carlo simulation the selected interval is reproduced. In the next step, the share of travellers in each interval was determined. On the other hand, using estimated Cox proportional duration model the probability of departure in each interval has been estimated.
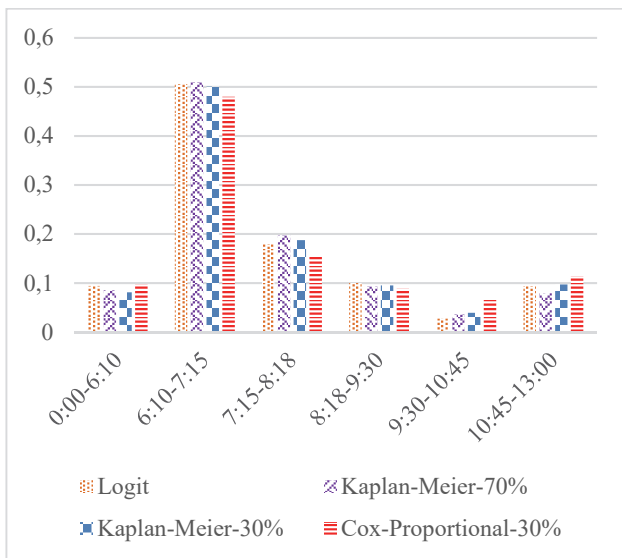

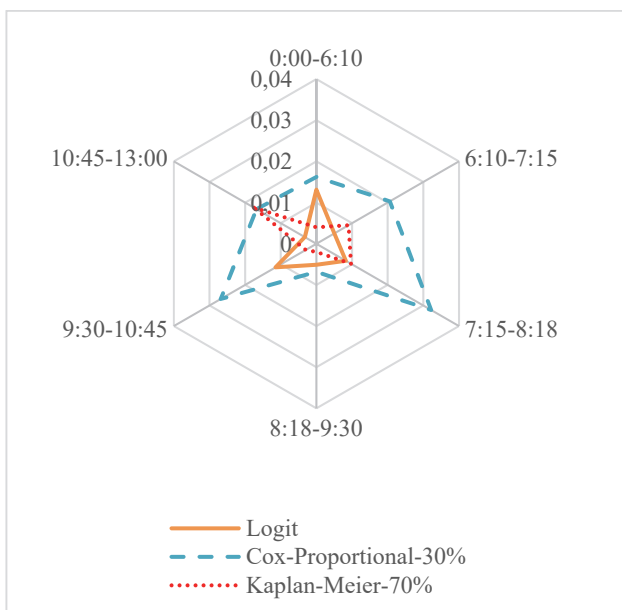**Figure 2** The probability of selection of each interval in various models

Figure 3 The difference between observed probability of selection each time intervals and its corresponding prediction

Also, non-parametric Kaplan-Meier analysis was done for predefined intervals and the results are shown in Fig. 2. As can be seen, in this figure the probability of departure time at each interval is presented. The blue column represents observed data (30% of data which have not been used in calibration) and models have to reproduce this column. It can be seen that Cox model underestimates for intervals between 6:15 and 9:30 although it overestimates for other intervals.

Fig. 3 shows the error of models prediction in a radar diagram and as can be seen, Cox model has more deviation from the observed value in comparison to others.

5 CONCLUSION

The results of this study show that the cluster-based MNL approach has predicted departure time for work trips accurately compared to the duration model in the case study of Mashhad city. In addition to predictability power, the sensitivity of this model to socioeconomic characteristics of travellers made it more reliable. Additionally, our use of six time periods is more disaggregate than a simple peak versus off-peak breakdown and therefore better serves the purposes of evaluating congestion management policies and air quality modelling.

The continuous modelling of departure time removes the necessity of discretization, which is inevitable in MNL, and it also provides fine temporal resolution. Nevertheless, this investigation shows that if time partitioning is done properly, the predictability power of MNL will be more than the duration model's predictability.

We also have to note this limitation that, rounding departure time to 5 or 15-minute periods by travellers in surveys resulted in partial violation of continuous time assumption in duration models.

As we expected, results of coefficient (parameter) estimation indicate traveller's job has a significant effect on departure time choice. This is predictable because of different start time and activity duration for various jobs. Moreover, the effect of selected mode on departure time choice is unsurprising. The public transportation users are more likely to depart in a period between 7:15 and 8:18. The pedestrians and bicycle riders are more reluctant to depart before 6:10 than the period between 6:10 and 10:45 and vice versa, they are more prone to leave their homes between 10:45 and 13:00. The long travel times will dictate choosing time periods before 6:10 for morning work shifts and period of 9:30 to 10:45 for afternoon shifts that are shown with positive coefficient in Tab. 1. The positive value of the coefficient for "parent" variable shows that parents depart earlier in order to accompany their children to school or day-care. Also, the positive value of "travel time" coefficient shows longer travel time increases the likelihood of sooner departure time due to the importance of on-time arrival for work trips. It can be seen Cox model underestimates for intervals between 6:15 and 9:30 although it overestimates for other intervals.

6 REFERENCES

[1] Saleh, W. & Farrell, S. (2005). Implications of congestion charging for departure time choice: Work and non-work

schedule flexibility. *Transportation Research Part A: Policy and Practice, 39*(7-9), 773-791. https://doi.org/10.1016/j.tra.2005.02.016

[2] Noland, R. B. & Small, K. A. (1995). Travel-time uncertainty, departure time choice, and the cost of the morning commute: Institute of Transportation Studies, University of California, Irvine.

[3] Bajwa, S., Bekhor, S., Kuwahara, M., & Chung, E. (2008). Discrete Choice Modeling of Combined Mode and Departure Time. *Transportmetrica, 4*(2), 155-177. https://doi.org/10.1080/18128600808685681

[4] Bhat, C. R. & Steed, J. L. (2002). A continuous-time model of departure time choice for urban shopping trips. *Transportation Research Part B: Methodological, 36*(3), 207-224. https://doi.org/10.1016/S0191-2615(00)00047-3

[5] He, S. Y. (2013). Does flexitime affect choice of departure time for morning home-based commuting trips? Evidence from two regions in California. *Transport Policy, 25*(0), 210-221. https://doi.org/10.1016/j.tranpol.2012.11.003

[6] Zhang, W.-Y., Guan, W., Sun, H.-J., & Mao, B.-H. (2015). Individual departure time decision considering departure scheduling utility. *Journal of Central South University, 22*, 787-792. https://doi.org/10.1007/s11771-015-2583-x

[7] Nurul Habib, K. M., Day, N., & Miller, E. J. (2009). An investigation of commuting trip timing and mode choice in the Greater Toronto Area: Application of a joint discrete-continuous model. *Transportation Research Part A: Policy and Practice, 43*(7), 639-653. https://doi.org/10.1016/j.tra.2009.05.001

[8] Biggiero, L., Cascetta, E., & Nuzzolo, A. (1992). Analysis and modelling of commuters' departure time and route choices in urban networks. *Sixth World Conference on Transportation Research*.

[9] Small, K. A. (1982). The scheduling of consumer activities: work trips. *The American Economic Review*, 467-479.

[10] Small, K. A. (1987). A discrete choice model for ordered alternatives. *Econometrica: Journal of the Econometric Society*, 409-424. https://doi.org/10.2307/1913243

[11] Hendrickson, C. & Plank, E. (1984). The flexibility of departure times for work trips. *Transportation Research Part A: General, 18*(1), 25-36. https://doi.org/10.1016/0191-2607(84)90091-8

[12] Afanadizadeh, S., Zahabi, S. A. H., & Kalantari, N. (2010). Forecasting the traffic share and environmental impacts of BRT using discrete choice models: case study of Isfahan. *8th International Congress on Civil Engineering*.

[13] Bhat, C. R. (1998). Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B: Methodological, 32*(6), 361-371. https://doi.org/10.1016/S0191-2615(98)00004-6

[14] Ettema, D. & Timmermans, H. (2003). Modelling departure time choice in the context of activity scheduling behaviour. *Transportation Research Record: Journal of the Transportation Research Board, 1831*(1), 39-46. https://doi.org/10.3141/1831-05

[15] Lim, Y. & Heydecker, B. (2005). Dynamic departure time and stochastic user equilibrium assignment. *Transportation Research Part B: Methodological, 39*(2), 97-118. https://doi.org/10.1016/j.trb.2003.08.003

[16] Ben-Elia, E., Bierlaire, M., & Ettema, D. (2010). A Behavioural Departure Time Choice Model with Latent Arrival Time Preference and Rewards for Peak-hour Avoidance. *Paper presented at the European Transport Conference*.

[17] Sasic, A. & Habib, K. N. (2013). Modelling departure time choices by a Heteroskedastic Generalized Logit (Het-GenL) model: An investigation on home-based commuting trips in the Greater Toronto and Hamilton Area (GTHA). *Transportation Research Part A: Policy and Practice, 50*(0), 15-32. https://doi.org/10.1016/j.tra.2013.01.028

[18] Afanadizadeh, S. & Safari, F. (2018). Using Clustering Methods in Multinomial Logit Model for Departure Time Choice.

[19] Hosmer, D. W. & Lemeshow, S. (1999). Applied Suvival Analysis. Regression Modelling of Time to Event Data.

[20] Afandizadeh, S., Khaksar, H., & Kalantari, N. (2013). Bus fleet optimization using genetic algorithm a case study of Mashhad. *International Journal of Civil Engineering, 11*(1), 43-52.

[21] Afandizadeh, S., Yadak, M., & Kalantari, N. (2011). Simultaneous determination of optimal toll locations and toll levels in cordon-based congestion pricing problem (case study of Mashhad city). *International Journal of Civil Engineering, 9*(1), 33.

[22] McMillan, T. E. (2007). The relative influence of urban form on a child's travel mode to school. *Transportation Research Part A: Policy and Practice, 41*(1), 69-79. https://doi.org/10.1016/j.tra.2006.05.011

[23] Greene, W. H. (2012). NLOGIT version 5.0: reference guide. Econometric Software, Inc, Plainview, NY.

[24] Schemper, M. & Stare, J. (1996). Explained variation in survival analysis. *Statistics in medicine, 15*(19), 1999-2012. https://doi.org/10.1002/(SICI)1097-0258(19961015)15:19<1999::AID-SIM353>3.0.CO;2-D

Contact information:

Shahriar Afandizadeh ZARGARI, Associate Professor
(Corresponding author)
Department of Transportation Engineering and Planning,
School of Civil Engineering, Iran University of Science & Technology,
P.O. Box 16765-163, Narmak, Tehran, Iran
E-mail: zargari@iust.ac.ir

Farshid SAFARI, PhD Candidate
Department of Transportation Engineering and Planning,
School of Civil Engineering, Iran University of Science & Technology,
P.O. Box 16765-163, Narmak, Tehran, Iran
E-mail: f_safari@iust.ac.ir

Appendix 1 Descriptive information of departure time for work trips

Duration	Hazard	Survival function	Hazard rate	Hazard rate S. E	Exited	At risk
00:00-4:30	0.0069	0.9931	0.0015	0.0002	60	8728
4:30-5:00	0.0097	0.9835	0.0195	0.0021	84	8668
5:00-5:30	0.0155	0.9683	0.0312	0.0027	133	8584
5:30-6:00	0.0549	0.9151	0.1129	0.0052	464	8451
6:00-6:30	0.1559	0.7725	0.3381	0.0095	1245	7987
6:30-7:00	0.4292	0.4409	1.0931	0.0195	2894	6742
7:00-7:30	0.3196	0.3000	0.7609	0.0213	1230	3848
7:30-8:00	0.2781	0.2165	0.6460	0.0236	728	2618
8:00-8:30	0.2307	0.1666	0.5215	0.0248	436	1890
8:30-9:00	0.2524	0.1245	0.5777	0.0298	367	1454
9:00-9:30	0.1766	0.1025	0.3875	0.0278	192	1087
9:30-10:00	0.1352	0.0887	0.2900	0.0263	121	895
10:00-10:30	0.0930	0.0804	0.1951	0.0230	72	774
10:30-11:00	0.1268	0.0702	0.2707	0.0286	89	702
11:00-11:30	0.5514	0.0315	1.5225	0.0766	338	613
11:30-12:00	0.9818	0.0006	3.8571	0.0622	270	275
12:00-12:30	0.6000	0.0002	1.7143	0.8942	3	5
12:30-13:00	1.0000	0.0000	4.0000	0.0000	2	2