Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

Modeling the risk perception of pedestrians using a nested logit structure

Babak Mirbaha[a]*, Atieh Asgari Toorzani[b]

[a] *Assistant Professor, Civil Engineering Department, Faculty of Engineering and Technology, Imam Khomeini International University, Qazvin, Iran*

[b] *PhD candidate, Civil Engineering Department, Faculty of Engineering and Technology, Imam Khomeini International University, Qazvin, Iran*

Abstract

Pedestrians are the most vulnerable road users since they do not have a protective shell. One of the most common collisions for them is pedestrian-vehicle at intersections. In order to develop appropriate countermeasures to improve safety for them, researches have to be conducted to identify the factors that affect the risk of getting involved in such collisions. More specifically, this study investigates factors such as the influence of walking alone or having a baby while crossing the street, the observable age of pedestrian, the speed of pedestrians and the speed of approaching vehicles on risk perception of pedestrians. A nested logit model was used for modeling the behavioral structure of pedestrians. The results show that the presence of more lanes at intersections and not being alone especially having a baby while crossing, decrease the probability of taking a risk among pedestrians. Also, it seems that teenagers show more risky behaviors in crossing the street in comparison to other age groups. Also, the speed of approaching vehicles was considered significant. The probability of risk taking among pedestrians decreases by increasing the speed of approaching vehicle in both the first and the second lanes of crossings.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: Pedestrians; Intersection; Nested logit; Risk.

1. Introduction

As a result of rapid urbanization, lack of adequate facilities, and lack of adherence to traffic regulations by all road users, accidents related to pedestrians have become a major safety problem around the world (Zhou et al. 2013). Road

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: mirbaha@eng.ikiu.ac.ir

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY

and traffic infrastructures in developing countries are neither suitable for pedestrians nor pedestrians' crossing behavior.

Different facilities such as pedestrian crossings (signalized and un-signalized), overpasses, and underpasses have been designed at intersections or mid-block crosswalks to help pedestrians' crossing. Pedestrian signals have been utilized as a means of safe crossing for all road users. However, in many cases, they won't be able to guarantee pedestrians' safe crossings. Pedestrian behavior is complex and easily influenced by environment and is strongly related to human factors. Accordingly, recognizing and analyzing their behaviors in crossings can improve the planning and design for their safety.

Regarding to previous studies, one of the most practical solutions of preventing pedestrians' crashes is identifying the crossing behavior of pedestrians including selected time to collision (TTC) or gap time (GT) and speed of approaching vehicle (Pawar and Patil, 2015, Koh and Wong, 2014, Onlecin and Alver, 2015). Risk perception while crossing the street requires identifying TTC and SAV.

According to Hayward (1972) Time-To-Collision (TTC) has proven to be an effective measure for rating the severity of conflicts. TTC defined as: "The time required for two vehicles to collide if they continue at their present speed and on the same path"(Hayward, 1972). Here TTC is defined as the time required for a vehicle and a pedestrian to collide.

Many studies have been conducted about pedestrian gap acceptance and its relationship with probability of being involved in crashes. For example, in Egypt, as a developing country, Serag (2014) examined pedestrians' crossing behaviors at un-signalized intersections by using video recording data using binary logit model and linear regression analysis. The results showed that the probability of pedestrian gap acceptance increases with the increase of the gap size, rolling gap, and frequency of attempt, whereas it reduces with an increase in vehicle speed.

Koh and Wong (2014) examined gap acceptance of violating pedestrians at signalized intersections using the binary logit model. According to their results, the majority (85%) of pedestrians reject gaps that are 6.3 s or less at near end crossings and 5.2 s or less at far end crossings. A binary logit model has been used by Pawar and Patil (2014) to investigate the temporal and spatial gap acceptance by pedestrians in mid-block crossings in India. The results showed that the probability of accepting a spatial gap decreases with an increase in speed. The accepted temporal or spatial gap size increases with an increase in the size of conflicting vehicle.

In another study, Pawar et al. (2016) analyzing and quantifying the dilemma zone for crossing pedestrians at high-speed uncontrolled midblock crossings. The dilemma zone can be used to develop a pedestrian assistance system at midblock crossing for pedestrian safety. Dilemma zone was determined by using different methods such as the gap cumulative distribution method, the binary logit method, the support vector machine and the probabilistic method. For the selected midblock sections, dilemma zone started at 49 m and ended at 62 m upstream from the marked pedestrian crossing.

Dutta and Vasudeyan (2017) conduct a research in the city of Kanpur, India to understand the effect of waiting time on pedestrian rolling gap acceptance behavior at un-signalized intersections in India under heterogeneous traffic. The results show male pedestrians behave more aggressively and undertake unsafe street crossings. When pedestrians are part of a group, they do not accept rolling gaps easily. They also found that when the approaching vehicle type is a two-wheeler or when vehicles are not in a platoon, the probability of accepting a rolling gap decreases.

Quistberg et al. (2015) estimate the risk of pedestrian collisions at intersections and mid-blocks in Seattle, WA. Their result in a multivariable model showed that intersections with 4 segments or 5 or more segments had higher pedestrian collision rates compared to mid-blocks. Non-residential roads, locations with traffic signals and locations that encourage greater pedestrian activity (more bus use, more fast food restaurants, higher employment, residential, and population densities) had significantly higher collision rates. Locations with a one-way road or those with signs encouraging motorists to cede the right-of-way to pedestrians had fewer pedestrian collisions.

Many studies studies have been conducted through gap acceptance but calibrating nested models for such a case is not common. Shankar et al. (1995) by using 5-year accident data from a 61 km section of rural interstate in Washington State (which has been selected as an ITS demonstration site), estimate a nested logit model of accident severity. Four levels of severity have been considered: (1) property damage only, (2) possible injury, (3) evident injury, and (4) disabling injury or fatality. The estimation results provide valuable evidence on the effect that environmental conditions, highway design, accident type, driver characteristics and vehicle attributes have on accident severity. Their findings show that the nested logit formulation is a promising approach to evaluate the impact that ITS or other safety-related countermeasures may have on accident severities.

Abdel-Aty and Abdelwahab (2003) modeled rear-end collisions using a nested logit structure. Four rear-end crash configurations are defined based on the type of the two involved vehicles (lead and following vehicles). An analysis of the effect of the geometric incompatibility of light truck vehicles (LTV) on drivers' visibility of other passenger cars involved in rear-end collisions has been presented. Results showed that driver's visibility and inattention in the following (striker) vehicle have the largest effect on being involved in a rear-end collision of configuration CarTrk (a regular passenger car striking an LTV).

As can be seen from this literature, the study of pedestrians' risk acceptance is not unexplored area of research but conducting a research to identify the effect of variables such as walking alone or having baby while crossing the street, the observable age of pedestrian in groups of teenagers, youth and elderly people, the speed of vehicles and the speed of pedestrians through a nested logit model has not been done yet.

In previous studies, we have demonstrated that factors such as individual characteristics (gender, age, dressing type, pedestrian speed), environmental conditions (other violating pedestrians, curb parking, waiting time) and traffic conditions (speed of approaching vehicles, TTC) can significantly affect pedestrians' risk-taking behavior (Jahandideh et al. 2017)

In this paper, based on the observed attitudes of pedestrians while crossing the street, two groups were identified which named adventurous and unadventurous. The main aspect of this division is the value of chosen TTC which is considered less than 3 seconds for adventurous group and more than 3 seconds for the latter. Nested logit models were calibrated to analyze the factors which are significant on the choice behaviors of pedestrians while crossing the street.

2. Body

2.1. Data

In this paper we are modeling the probabilities of four alternatives in two-risk intensity of risk acceptance based on the speed of approaching vehicle (SAV) and critical time to collision (TTC). In this regard, recorded video of the crossing pedestrians' behavior has been collected by the installed cameras in intersections. The quality of pedestrians' behavior has been investigated by processing images in ImageJ software and calibrating the videos based on field measurements.

The study was conducted in Qazvin which is the largest city and capital of the Province of Qazvin in Iran and Located in 150 km northwest of Tehran. At the 2015 census, its population was 596,932. Six intersections which have crowded pedestrian crossings with different widths and without countdown signals have been selected in the city. The study locations were geographically spread across the city. Each intersection characteristics is presented in Table 1 and Figure 1 shows the locations.

Pedestrians' crossing has been collected during peak hours (8–9:30 am and 4–6 pm) by mounted video cameras. Video coverage included pedestrians' waiting areas at each end of the pedestrian crossings, the length of crosswalk, and the traffic light. The recording days were three consecutive working days in the October 2015. Each video recording was continued for at least 90 minutes without interruption. TTC of 800 pedestrians (400 samples in signalized and 400 samples in un-signalized intersections) have been recorded.

Different lengths, including the crosswalk and the route, were measured through a field measurement for further analysis. After collecting data, different variables such as speed of vehicles, pedestrians' speed, and TTC were calculated by using the ImageJ software and calibrating the videos and images. It is worth noting that time parameter is accessible via the software.

Table1. Data description

Intersection	Number of lane	Length of crosswalk (m)	Average cycle length (sec)	Average green time (sec)	Average pedestrian volume (every 15 min)	Average vehicle volume (every 15 min)
Adl	3	18.6	90	50	154	230

Valiasr	3	21.3	130	30	123	488
Ferdosi	2	11.7	40	20	73	194
Shohada	2	7	100	30	50	234
Keyhan	2	12.5	Un-signalized		165	368
Valiasr mid-block	2	10	Un-signalized		59	363

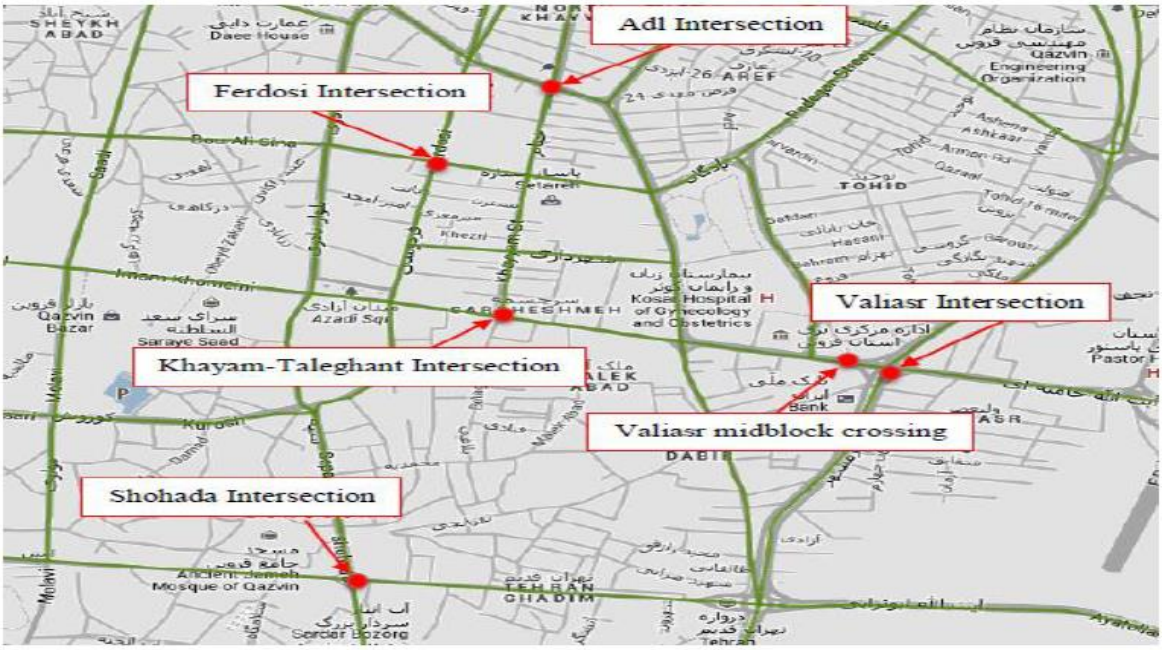

Fig. 1. Selected intersection of the study

2.2. Methodology

The nested logit model requires a priori specification of homogenous sets of alternatives (i.e. partitioning the set J into several exclusive groups). To fix the idea of a nested logit model, suppose that J alternatives can be divided into G mutually exclusive subgroups. Logically, one may think of the choice process as that of choosing among G choice sets and then making the specific choice j within the chosen set g, g = 1, . . . , G. Then, the mathematical form for a two-level nested logit model is as follows (Greene, 2003) the set J into several exclusive groups).

$$P_n(j,g) = P_n(g)P_n(j \mid g), \quad j = 1, 2, \ldots, j-1 \tag{1}$$

$$P_n(j \mid g) = \frac{\exp(\beta_j X_n \mid g)}{\sum_{j\epsilon g} \exp(\beta_j X_n \mid g)} \tag{2}$$

$$P_n(g) = \frac{\exp(\gamma_g M_g + \tau_g I_g)}{\sum_{g=1}^{G} \exp(\gamma_g M_g + \tau_g I_g)} \tag{3}$$

$$P_n(J) = 1 - \sum_{j=1}^{J-1} P_n(j) \tag{4}$$

Where $P_n(j, g)$ is the probability of a subject n to belong to category j in group g, Pn(j|g) the conditional probability of a subject n to belong to category j given that subject belongs to group g, Pn(g) the probability of subject n to belong to group g, Xn a vector of measurable attributes, Mg a vector of groups' characteristics, $Ig = \ln\_\sum_{j \epsilon g} \exp\left(\beta_j X_n \mid g \right)$ and is called the inclusive value of group g, βj and γg are vectors of coefficients to be estimated using the maximum likelihood approach, and τg is the coefficient of the inclusive value of group g. If we restrict all inclusive value parameters to one, then the nested logit model will be reduced to the MNL. The nested logit model can be easily extended to higher levels (more than two levels). The complexity of the model increases geometrically with the number of levels. But the model has been found to be extremely flexible and is widely used in transportation modeling (Abdel-Aty and Abdelwahab, 2003).

Goodness of fit for evaluated model (a value between zero and one) was defined based on equation 5(Kanafani, 1983).

$$\rho^2 = 1 - \frac{LL_\beta}{LL_0} \tag{5}$$

As illustrated in Hensher et al. (2005) for all nested logit models, there will exist a unique Inclusive Value (IV) parameter for each trunk, limb, and branch specified as part of the tree structure. As with other parameters, the analyst may constrain several of these IV parameters or normalize2 them (i.e. fix them) to equal some particular value (usually, but not always 1.0).

2.3. Results

The aim of the study is to consider pedestrians' risk acceptance. As can be seen in Fig. 2 the final nested logit model which has been considered has two nests, each of which has two risk intensity alternatives.

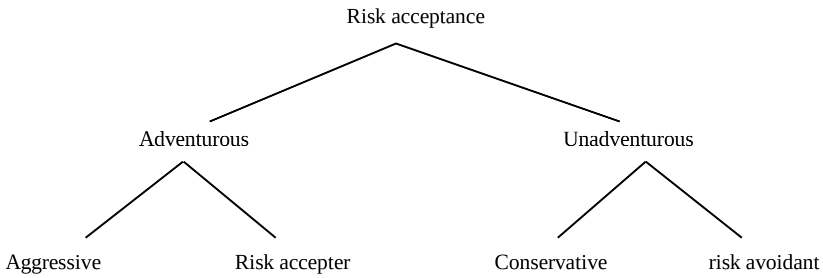

Fig. 2. Configuration of nested logit model

The definition of variables can be observed in Table 2 and the calibrated nested logit model is shown in Table 3. As can be seen, the significant variables of the calibrated nested logit model by NLOGIT 4.0 include Gap1, Gap2, SVeh1, SVeh2, Lane, Sped, Alone, Run, Baby, Teen, Suit.

Table 2. Definition of variables of nested logit model for risk

Category	Variable	Definition	Type
Individual characteristics	Sped	Speed of pedestrian	Continuous
	Run	Running pedestrian (At speed of more than 1.9 m/s)	Binary
	Teen	Perceived age of pedestrian (Teenager)	Binary
	Suit	Pedestrian dressing type (who wore suit)	Binary
	Baby	Pedestrians who accompanied a baby while crossing	Binary
	Alone	Pedestrians who were alone while crossing	Binary

Environmental characteristics	Gap1	Time to collision in the first lane of crossing	Continuous
	Gap2	Time to collision in the second lane of crossing	Continuous
	SVeh1	Speed of approaching vehicle in the first lane of crossing	Continuous
	SVeh2	Speed of approaching vehicle in the second lane of crossing	Continuous
Traffic condition	Lane	Number of lane	Continuous

The definition of variables can be observed in Table 2 and the calibrated nested logit model is shown in Table 3. As can be seen, the significant variables of the calibrated nested logit model include Gap1, Gap2, SVeh1, SVeh2, Lane, Sped, Alone, Run, Baby, Teen, Suit.

Table 3. Calibrated nested logit model for risk

Risk	Alternative	Variable	Coefficient	Standard Error	t-statistics
Adventurous	Aggressive	Gap 1	-0.46755***	0.07178	-6.51
		SVeh 1	0.04948**	0.02224	2.22
		SVeh 2	.19978***	0.02700	7.40
		Lane	-1.29579***	0.34686	-3.74
	Risk Acceptant	Baby	-2.41121**	1.02655	-2.35
		Run	-2.30135***	0.51456	-4.47
		Teen	1.54344***	0.51364	3.00
		Suit	1.60368***	0.35884	4.47
		Sped	3.19574***	0.34961	9.14
		Gap 1	-0.43066***	0.6432	-6.70
Unadventurous	Conservative	Constant	-6.67199***	0.92007	-7.25
		Gap 2	.44414***	0.11455	3.88
		SVeh 1	0.13658***	0.01936	7.05
		SVeh 2	0.17301***	0.02125	8.14
		Sped	-1.77255***	0.49927	-3.55
	Risk Avoidant	Alone	-0.64182**	0.22076	-2.91
		Run	-2.07935***	0.60877	-3.42
		Gap 2	0.49184***	0.11941	4.12
IV Parameters					
Adventurous			0.5(fixed)	Fixed parameter	3.74
Unadventurous			0.66379***	0.14560	4.07
McFadden Pseudo R-squared				0.402	

Note: ***, **, * ==> Significance at 1%, 5%, 10% level.

As can be seen in the calibrated model, increasing the time to collision in the first lane of crossing of intersection decreases the probability of risking by pedestrians. This can be interpreted from the negative coefficients of Gap1 in both aggressive and risk acceptant alternative of adventurous branch.

Aggressive pedestrians have more intention to risk with increasing the speed of approaching vehicle in both the first and the second lanes of crossing. The presence of more lanes at the intersection will lead to a reduction in the likelihood of aggressive behavior among pedestrians. This means that pedestrians tend to risk less at intersections with more lanes.

Tendency of risking among teenagers are much more than other age groups, which put them in risk acceptant group. On the contrary, pedestrians with baby do not tend to risk while crossing the street that could be interpreted from the negative coefficient of its variable in risk acceptant group. It is observed in the unadventurous branch that increasing the time to collision in the second lane of crossing increases the probability of risking.

Pedestrians' speed is an important factor in the risk acceptance. Increasing the pedestrians' speed led to decrease the probability of being conservative and on the contrary, increases the probability of being in the risk acceptant group.

If pedestrians are not alone while crossing the intersection, they will not risk. This can be interpreted by the negative coefficient of Alone in the risk avoidant group. Also, increasing the time to collision in the second lane of crossing of intersection decreases the probability of risking by pedestrians. This can be interpreted from the positive coefficients of Gap2 in both conservative and risk avoidant groups.

The probability of not risking increases by increasing the speed of approaching vehicle in both the first and the second lanes of crossing. This can be observed from the positive coefficient of the speed of approaching vehicle in both the first and the second lanes of crossing in the conservative group.

Pseudo-R^2 is a measure that has the same kind of interpretation of R^2 in the linear model. As Mcfadden illustrated, Pseudo-R^2 between 0.2 and 0.4 is comparable with 0.7 and 0.9 for R^2 in the linear model (Hensher et al., 2005).

The alternative finding of a significant IV parameter estimate suggests that the parameter is not equal to zero, but does not indicate whether the parameter lies outside the upper bound of the 0–1 range (recall that an IV parameter cannot be less than zero). Thus, for significant IV parameters, a second test is required to determine whether the upper bound has been exceeded. This test may be undertaken with a simple modification to the test conducted to determine whether the parameter is statistically equal to zero. This modification is shown by wald-test which is a test of whether a Wald-statistic is significantly different to zero and wald-statistics is the ratio of an importance weight to its standard error (Hensher et al., 2005).

$$Wald - test = \frac{IV_{parameter} - 1}{Standard\ Error} = \frac{0.66379 - 1}{0.14560} = -2.309 \qquad (6)$$

Table 4 summarizes direct elasticity of the variables included in the calibrated nested logit model. This table shows estimates of the effect of the explanatory variables included in the model on the probability of a certain risk configuration while taking into account the impact of the other explanatory factors. For example, number of lane has the largest negative effect of having a risk of configuration adventurous (aggressive). This emphasizes that with increasing the number of lanes at intersection, the aggressive behavior decreases among pedestrians.

Pedestrians who accompanied a baby while crossing the intersection, has the largest negative impact on adventurous (risk acceptant) group. This also shows that the risk acceptance decreases among pedestrians who cross with babies.

In terms of the speed of approaching vehicle in the both first and second crossing lane, the aggressive and conservative behaviors are both increases. This might be due to pedestrians' wrong judgment about the speed of approaching vehicles. Also, time to collision in the second lane of crossing has the largest elasticity in the risk avoidant group which seems to be reasonable.

Table 4. Elasticity of the calibrated nested logit model for risk

Variable	Direct Elasticity
Gap 1 specific to Adventurous (Aggressive)	-1.28
Gap 1 specific to Adventurous (Risk acceptant)	-0.54
Gap 2 specific to Unadventurous (Conservative)	1.6
Gap 2 specific to Unadventurous (Risk avoidant)	0.86
SVeh 1 specific to Adventurous (Aggressive)	0.72
SVeh 1 specific to Unadventurous (Conservative)	1.87
SVeh 2 specific to Adventurous (Aggressive)	2.54
SVeh 2 specific to Unadventurous (Conservative)	2.57
Lane specific to Adventurous (Aggressive)	-2.75
Baby specific to Adventurous (Risk acceptant)	-0.71
Run specific to Adventurous (Risk acceptant)	-0.045

Run specific to Unadventurous (Risk avoidant)	-0.071
Suit specific to Adventurous (Risk acceptant)	0.03
SPed specific to Adventurous (Risk acceptant)	0.78
Sped specific to Unadventurous (Conservative)	-1.54
Alone specific to Unadventurous (Risk avoidant)	-0.12

3. Conclusion

This study has been conducted to identify variables affecting pedestrians' intensity via calibrating a nested logit model. An observational survey has been conducted at six intersections located in Qazvin city and TTC of 800 pedestrians (400 samples in signalized and 400 samples in un-signalized intersections) have been recorded. Using a cumulative percentage graph of recorded TTCs, a ranking pattern for pedestrian risk condition has been proposed.

The presence of more lanes at intersections and not being alone especially crossing the street while accompany a baby decrease the probability of risking among pedestrians. Also, it seems that teenagers are more risky in crossing the street in comparison to other age groups.

Theoretically, it might be considered that pedestrians won't have a proper judgment on the speed of approaching vehicle but practically this study shows that they will. And the probability of risking decreases by increasing the speed of approaching vehicle in both the first and the second lanes of crossing.

The findings are expected to provide a relative cognition on pedestrians' behavior but not a comprehensive one, because pedestrians' preferences should be explained by discrete choice models. As an instance, the effect of socioeconomic characteristics of pedestrians has not been investigated in this study due to time deficiency and burden of census. It would be worthwhile that a comprehensive research is conducted which is based on a questionnaire and include demographic characteristics of pedestrians.

4. References

Abdel-Aty M., Abdelwahab H., 2003, Configuration Analysis of Two-Vehicle Rear-End Crashes, Transportation Research Record: Journal of the Transportation Research Board 1840, 140-147.

Dutta B., Vasudevan V., 2017, Study on Pedestrian Risk Exposure at Unsignalized Intersection in a Country with Extreme Vehicle Heterogeneity and Poor Lane Discipline, Transportation Research Record: Journal of the Transportation Research Board Volume 2634: 69–77.

Greene, William H., 2003, Econometric Analysis. New Jersey: Pearson Education, Inc., Upper Saddle River.

Hayward, J.Ch., 1972, Near miss determination through use of a scale of danger. report no TTSC 7115, Pennsylvania: The Pennsylvania State University.

Hensher, D. A., Rose, J. M., Greene W. H., 2005, Applied choice analysis. New York: cambridge university press.

Jahandideh Z., Mirbaha B., Rassafi A. A., 2017, Modeling the risk intensity of crossing pedestrians in intersections based on selected critical time to collision (a case study of Qazvin city), Transportation research board 96th annual meeting, Washington DC, USA.

Jahandideh Z., Mirbaha B., Rassafi A. A., 2017, Identifying factors affecting pedestrians' crossing decisions at intersections in Iran, Proceeding of institution of civil engineers-municipal engineer.

Kanafani A., 1983, Transportation Demand Analysis, McGraw-Hill, Inc., New York, NY, USA.

Koh, P. P. and Wong, Y. D., 2014, Gap acceptance of violators at signalised pedestrian crossings, Accident Analysis and Prevention 62, 178-185.

Onelcin, P. and Alver, Y., 2015, Illegal crossing behavior of pedestrians at signalized intersections: factors affecting the gap acceptance, Transportation research part F: traffic psychology and behavior, 31, pp.124-132.

Pawar, D. S., Patil G. R., 2015, Pedestrian temporal and spatial gap acceptance at mid-block street crossing in developing world, Journal of Safety Research, 39-46.

Pawar, D. S., Kumar V., Singh N., and Patil G. R., 2016, Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials, Safety Science, 295-303.

Quistberg, A. D., Beth, E. J., Ebel, E., Moudon, A., Saelens, E., Hurvitz, M. P., 2015, Multilevel models for evaluating the risk of pedestrian–motor vehicle collisions at intersections and mid-blocks, Accident Analysis & Prevention, 99-111.

Serag, M. S., 2014, Modeling pedestrian road crossing at uncontrolled mid-block locations in developing countries, Int J Civ Struct Eng 4(3), 274-286.

Shankar, V., Mannering, F., Barfield, W., 1995, Effect of roadway geometrics and environmental factors on rural accident frequencies, Accident Analysis and Prevention, 371-389.

Zhou Z. P., Liu Y. S., Wang,W., Zhang Y., 2013, Multinomial logit model of pedestrian crossing behaviors at signalized intersections, Discrete Dynamics in Nature and Society.