ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

AN ANALYSIS OF PEDESTRIAN WAITING TIME AT UNCONTROLLED CROSSWALKS USING DISCRETE CHOICE MODEL

M.Sc. THESIS

Ehsan AMIRNAZMIAFSHAR

Department of Civil Engineering

Transportation Engineering Programme

December 2018

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

AN ANALYSIS OF PEDESTRIAN WAITING TIME AT UNCONTROLLED CROSSWALKS USING DISCRETE CHOICE MODEL

M.Sc. THESIS

Ehsan AMIRNAZMIAFSHAR
(501151425)

Department of Civil Engineering

Transportation Engineering Programme

Thesis Advisor: Assoc.Prof. Dr. Hüseyin Onur Tezcan

December 2018

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

AYRIK SEÇİM MODELİ KULLANILARAK KONTROLSÜZ YAYA GEÇİTLERİNDE YAYA BEKLEME SÜRESİNİN İNCELENMESİ

YÜKSEK LİSANS TEZİ

Ehsan Amirnazmiafshar
(501151425)

İnşaat Mühendisliği Anabilim Dalı

Ulaştırma Mühendisliği Programı

Tez Danışmanı: Assoc.Prof. Dr. Hüseyin Onur Tezcan

Aralık 2018

Ehsan Amirnazmiafshar, a M.Sc. student of İTU Graduate School of Science Engineering and Technology student ID 501151425, successfully defended the thesis entitled "AN ANALYSES OF PEDESTRIAN WAITING TIME AT UNCONTROLLED CROSSWALKS USING DISCRETE CHOICE MODEL", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor: **Assoc. Prof. Dr. Hüseyin Onur TEZCAN** ....
Istanbul Technical University

Jury Members: **Prof. Dr. Kemal Selçuk ÖĞÜT** ....
Istanbul Technical University

Prof. Dr. Yetiş Şazi MURAT ....
Pamukkale University

Date of Submission : 17 September 2018
Date of Defense : 4 December 2018

FOREWORD

As a candidate of transportation engineer, I aim to improve the transportation systems. My dear advisor Assoc. Prof. Dr. Hüseyin Onur Tezcan encouraged me to focus on pedestrians waiting time analysis to reach the purpose. This topic is related to transportation planning. Assoc. Prof. Dr. Hüseyin Onur Tezcan by giving much motivation and helping me step by step patiently in all this time has played a significant role in completing this thesis. I truly appreciate all he has done for me in the last two years. I would also like to thank Prof. Dr. Kemal Selçuk ÖĞÜT for his help and advice which have been invaluable to me through all stages of my thesis.

December 2018

Ehsan AMIRNAZMIAFSHAR
Civil Engineer

TABLE OF CONTENTS

Page

ABBREVIATIONS

s	**:** Second
m/s	**:** Meter per second
e	**:** Exponential function
m	**:** Meter
km/h	**:** kilometer per hour
VPH	**:** Vehicles per hour
PWS	**:** Number of pedestrians start waiting at the same time
LOS	**:** Level of service

LIST OF TABLES

Page

LIST OF FIGURES

Page

AN ANALYSIS OF PEDESTRIAN WAITING TIME AT UNCONTROLLED CROSSWALKS USING DISCRETE CHOICE MODEL

SUMMARY

Pedestrians are vulnerable road users of the traffic system. An uncontrolled mid-block crosswalk is a conflict area between pedestrians and vehicles where is the high-risk area for pedestrians. Pedestrian crossing behavior is analyzed to improve their safety while crossing the street and the provision of appropriate pedestrian facilities in the area. A study of pedestrians crossing behavior is conducted at an uncontrolled mid-block crosswalk in Istanbul-Turkey to model the pedestrians waiting time which is related to their behavior for making the crossing decision. This thesis focused on the issues encountered in the modeling of the operational behavior of pedestrians. Data are gathered by video recording from the pedestrians who did not know that the camera was recording their behavior.

A comprehensive literature review is conducted for various parameters of pedestrians crossing behavior to decide for crossing the street. Most of these parameters and some additional parameters which are about the traffic characteristics and pedestrians characteristics are studied. Pedestrians waiting time and other related factors to their waiting time are extracted from the video recording of 618 pedestrians behaviors and the traffic stream. The results of both the Kruskal-Wallis H test and Mann-Whitney test show that the pedestrians age can be classified into two groups, which do not come from the same population includes the elderly or young pedestrians. Also, the result of these tests presents that gender can be classified into two groups including males and females. The number of rejected vehicles by each pedestrian and the number of pedestrians start waiting at the same time at the curbside to cross the street is counted. Moreover, the average headway (the time between the front bumper of one vehicle and the front bumper of the next) of rejected vehicles by pedestrians is determined.

The discrete choice framework is used because of its capacity to deal with individuals' choice behavior. Pedestrians waiting time is classified into three levels that include

low, medium and high levels based on the level of service of pedestrians waiting time. The results of the model show that males do not prefer to have a high level of waiting time compared to females. Also, the probability of waiting time to be at the high level for the elderly pedestrians is higher than low level and the probability of being at the medium level of waiting time is more than being at the high level. The probability of waiting time to be at the high level is less than low and medium levels if the number of pedestrians who start waiting at the same time at the curbside to cross the street increases. Also, the higher the number of rejected vehicles by the pedestrians, the higher the probability of waiting time to be at a high level compared to the low and medium level. Furthermore, it is found that as much as the average of rejected vehicles which in this study, is assumed equal to 0 for pedestrians whose waiting time is 0 and for other pedestrians is from 1.25s to 5.01s is higher for the the pedestrian, the probability of waiting time at the high level increases compared to the low and medium level..

By knowing the results of this study, the urban transport planners and engineers can have a better understanding of pedestrians behavior, create proper design guidelines and inform policy decisions.

AYRIK SEÇİM MODELİ KULLANILARAK KONTROLSÜZ YAYA GEÇİTLERİNDE YAYA BEKLEME SÜRESİNİN İNCELENMESİ

ÖZET

Yayalar trafik düzeninin savunmasız yol kullanıcılarıdır. Kontrolsüz yaya geçidi, yayalar için yüksek risk taşır ve aynı zamanda yayalar ve taşıtlar arasındaki bir kesişim bölgesidir. Yaya geçidindeki yaya hareketleri, caddeyi geçerken ve bölgedeki uygun yaya olanaklarının sağlanması ve güvenliğinin arttırılması amacıyla analiz edilir. Yayaların geçişi ile ilgili olarak yayaların bekleme sürelerini modellemek için İstanbul-Türkiye'de kontrolsüz bir yaya geçidinde yaya geçişi davranışları incelenmiştir.

Bu tez, yayaların halihazırda davranışlarının modellenmesinde karşılaşılan sorunlara odaklanmıştır. Veriler, İstanbul'da yer alan Aytar Caddesi'ndeki kontrolsüz bir yaya geçidinde, davranışlarının kaydedildiğini bilmeyen yayaları çeken kamera tarafından kaydedildi. Kayıt, normal bir iş gününde on iki saat boyunca yapıldı. Kontrolsüz yaya geçidindeki yayaların ve trafik özelliklerinin, yayaların geçiş davranışlarını analiz ettiği görülmektedir.

Caddeyi geçmeye karar vermek ile ilgili yaya geçiş davranışlarının çeşitli parametreleri için kapsamlı bir literatür taraması gerçekleştirilmiştir. Araştırma kapsamında literatürde karşılaşılan parametrelerin çoğu, trafiğin özellikleri ve yayaların özellikleri olan bazı ek parametreler de incelenmiştir. Yaya bekleme süresi değişkeni bağımlı ile beş bağımsız değişken analiz edilerek 618 yayanın davranışları incelenmiştir.Yayalar cinsiyetleri açısından Kruskal Testi'nde ve Mann-Whitney U Test'te değerlendirildiklerinde, erkeklerin ve kadınların aynı toplumdan olmadığı belirlenmiştir. Yayaların yaşı bu testlerde iki gruba ayrılarak incelenmiş ve doğruluğu test tarafından onaylanmıştır. Bu gruplardan biri aynı toplumdan gelmeyen (aynı tip yaya davranışında bulunmayan) yaşlı olmayan ve diğer grup ise yaşlı yayaları içermektedir. 618 yayanın özellikleri gözlemlendiğinde, yayaların cinsiyetlerinin

yaşlarının bekleme süresiyle ilişkili olduğu belirlenmiştir. Bu gözlemde 359 yaya erkek, 259 yaya kadın olarak bulunmaktadır. Yayaların yaşı, görünümleriyle doğru orantılı olarak tahmin edilmiştir. Cinsiyet, yaş parametrelerinin yanı sıra yayaların bekleme süresince cep telefonu kullanıp kullanmama durumları da değerlendirilmiştir. Kruskal testine göre bekleme süresince cep telefonu kullanan ve kullanmayan yayalar aynı toplumda (aynı tip yaya davranışı gösteren) bulunmaktadır. Bu nedenle cep telefonu kullanma durumu ayrık seçim modeli içerisinde değerlendirilmemiştir. Diğer değişkenler ise sürücünün yaya geçidinde bekleyen yaya gördüğünde hızını azaltıp azalmama durumu, yayaların yaya geçidinden geçerken izlediği yolun düz ya da çapraz şekilde olması durumu, yaya geçidinden geçerken yayaların elinde bir şeyler taşıması ya da taşımaması durumu ve yaya geçidinin bulunduğu alanda park halinde olan araçların bulunup bulunmaması durumudur. Bu değişkenler Kruskal Testi'nde ve Mann-Whitney U Test'te değerlendirildiğinde, modelde değişken olarak kullanılamayacağı tespit edilmiştir.

Yayaların yaya geçidinde beklerken, karşıdan karşıya geçmeye karar verme süresinde yaya geçidini henüz geçmiş taşıtın arkası ile yaya geçidini henüz geçmemiş taşıtın ön kısmı arasında kalan zamana boşluk denmektedir. Söz konusu bu değişken modelde kullanılmamıştır. Çünkü yayaların bekleme süresi ile arasında bulunan kolerasyon pozitif olarak bulunmuştur. Bulunan bu pozitif kolerasyon beklenmeyen bir sonuçtur.

Yapılan gözlemler sırasında belirlenmiş olan unsurlardan biri de yaya geçidinden geçmek için aynı anda beklemeye başlayan yaya sayısıdır. Bu parametrenin kolerasyonu modelde kullanmak için uygun olduğundan, modelde bir unsur olarak kullanılmıştır.Yayaların yaya geçidinde beklerken, karşıdan karşıya geçmeye karar verme süresinde yaya geçidini henüz geçmiş taşıtın önü ile yaya geçidini henüz geçmemiş taşıtın ön kısmı arasında kalan zamana headway denmektedir. Yaya geçidini geçmek için bekleyen yayaların, yaya geçidinde beklediği sürede geçen taşıtların headwaylarinin ortalama değeri de ayrı bir unsur olarak modelde kullanılmıştır. Yaya geçidinde bekleme süresi sıfır olan yayalar için bu unsur sıfır olarak değerlendirilmiştir. Modelde kulllanılmış olan diğer bir unsur da yaya geçidinde bekleyen yayaların, beklediği süre içerisinde geçen araç sayısıdır. Bekleme süresi sıfır olan yayaların bu unsur değerleri de sıfırdır.

Ayrık seçim modeli bireylerin seçim davranışlarının değerlendirilmesi amacıyla kullanılmıştır. Yayalar bekleme süreleri düşük, orta ve yüksek seviyeleri içeren üç seviyeye ayrılır. Ayrık bir seçim modeli kullanmak için, sürekli veri olan bekleme süresi ayrı verilere dönüştürülür. Veriler, her biri için eşik değerleri belirleyerek üç seviyeye ayrılır. Bu üç seviye, yayaların bekleme süresi için hizmet seviyesi olan A, B ve C'den F'e kadar olan düşük, orta ve yüksek bekleme süreleridir.

Modelin sonuçları, erkeklerin kadınlara göre daha fazla süre beklemeyi tercih etmediğini göstermektedir. Yaşlı insanların bekleme sürelerinin yüksek seviyede olma ihtimali düşük seviyede olma ihtimalindenden daha yüksektir. Yine yaşlı insanların bekleme sürelerinin orta seviyede olma ihtimali yüksek seviyede olma ihtimalinden daha yüksektir. Aynı anda yaya geçidinde bekleyen yaya sayısı arttıkça bekleme süresinin düşük ve orta seviyede olması itimali yüksek seviyede olması ihtimalinden daha yüksek olur. Yaya geçidinde yaya beklerken geçen taşıtların headwaylerinin ortalaması ne kadar yüksek ise yaya geçidinde yaya bekleme süresi seviyesinin yüksek olma ihitmali düşük ve orta olması ihtimalinden daha düşüktür. Diğer bir sonuç ise yaya geçidinde yaya beklerken geçen araç sayısı ne kadar yüksek ise yaya geçidinde yayaların bekleme süresinin yüksek olması ihtimali orta ve düşük olması ihtimalinden daha yüksektir.

Bu çalışmanın sonuçlarının yukarıda bahsedilen kısımları bilindiğinden, ulaşım planlamacıları ve mühendisleri yayaların davranışlarını daha iyi anlayabilir, uygun tasarım ilkeleri oluşturabilir ve şehir planlamada yaya ulaşımı için gerekli stratejiler oluşturmada bilgi verebilirler.

1. INTRODUCTION

In recent decades, new rules and policies for pedestrians walking have been set up in many countries to consider it as a transportation mode. By this way, the pedestrians motivation to walk might be increased. These rules and policies can decrease the use of private vehicles, and it helps sought to reduce traffic and pollution in cities. Moreover, increasing pedestrian walking might reduce obesity in the communities that leads to increasing people's health.

The pedestrian has the least delay compared to other modes. However, in the metropolitan cities of developing countries, they received low levels of priority. Especially, as the number of trips made by motorized vehicles increases, the risk of the involvement of pedestrians in accidents increases. Therefore, they are vulnerable users of the traffic system (Keegan and O'Mahony, 2003; Tiwari et al., 2007; Ishaque and Noland, 2008). In 2010, pedestrians accounted for 21.9% of 1850 fatalities on roadways in the UK, and 46% of 126 deaths in road accidents were pedestrians in London. Globally, 22% of road fatalities were pedestrians in 2010 (World Health Organization, 2013).

The pedestrian behavior is examined in different fields such as urban planning, architecture, land use and even marketing (Okazaki and Matsushita, 1993; Parker et al., 2003). The studies on pedestrians behavior are related to the analysis of perceptual, attitudinal, psychological and motivational factors; all of them are factors related to different attributes of human beings (Bernhoft and Carstensen, 2008; Moyano Díaz, 2002; Evans and Norman, 1998). Eventually, analyzing the pedestrians behavior can help reduce the number of accidents that involve pedestrians in urban areas and increase their safety level (Lassarre et al., 2007).

Studying on the pedestrians waiting time which is directly related to both the pedestrian decision for starting to cross and their safety while the crossing the road is a significant issue. Pedestrians detect the traffic; consider different factors, which affect their decision to cross the street. (Lee et al. , 1984; Oudejans et al., 1996).

In this thesis, the waiting time of pedestrians is analyzed by discrete choice analysis to determine the effect of different factors on pedestrians crossing behavior accurately. The study area is a mid-block crosswalk at an uncontrolled location. The uncontrolled crosswalk is an interaction area between pedestrians and vehicles that can be a dangerous place and threaten their safety because of the high probability of accidents that can occur at this type of crosswalk. Also, traffic control devices that can improve safety are not in place or operation to dictate pedestrian movements at these types of crosswalks.

This thesis consists of five chapters. The first chapter is the introduction, which includes motivation, objectives, and organization of the thesis. In Chapter 2, a literature study is presented. In Chapter 3, some information about the data, data collection approach and analysis of data are presented. In Chapter 4, the model literature, which includes some information about the method, is given. Also the estimated model, its results and discussions about the results are presented. Finally, the conclusion and a summary of the findings are given in Chapter 5.

2. LITERATURE REVIEW

When the pedestrians arrive at a crosswalk, there are different factors, which affect their decision for crossing the street. They check the traffic stream to find an acceptable gap between vehicles to cross. When pedestrians reject more gaps, their waiting time increases at the curbside. Two major crossing patterns of pedestrians are one-step crossing (pedestrians cross the street without waiting near the median) and two-step crossing (pedestrians cross up to the median in one go and subsequently cross the far -side). Far-side is the other side of the street. When pedestrians cross the street in a straight path, it is called perpendicular crossing, while oblique crossing is the crossing in which pedestrians cross the street in a zig-zag manner. The mixed crossing is the combination of both oblique and perpendicular types of crossing (i.e., pedestrian crosses first half of the street in a straight path up to the median and the next half the pedestrian crosses in a zigzag manner and vice versa) (Asaithambi et al., 2016). In Figure 2.1, this classification is depicted. Figure 2.1a presents the classification of crossing according to the steps, and Figure 2.1b shows the path in part a and b, respectively.

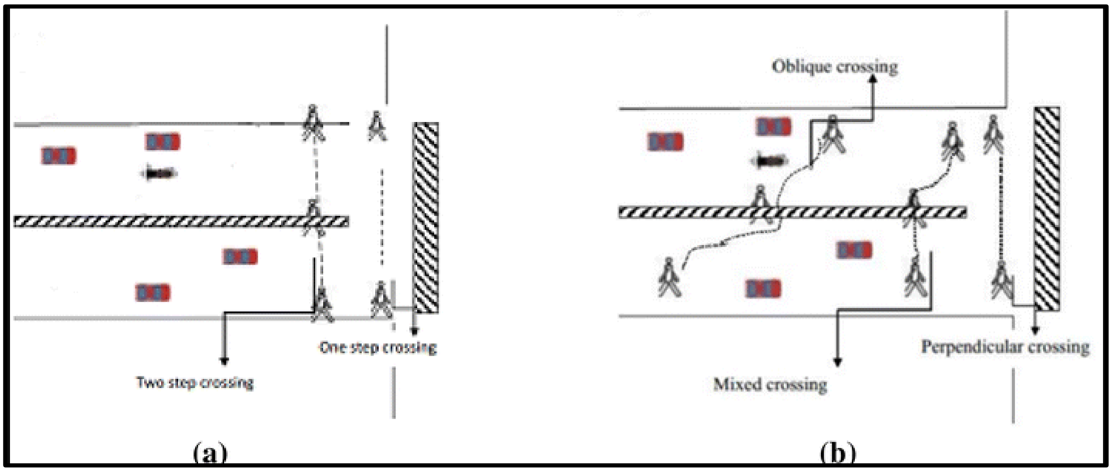

Figure 2.1: Classification of crossing **(**a) according to the steps (b) according to the path (Asaithambi et al., 2016).

The crosswalks are classified into two types as at-grade crosswalks and grade separated crosswalks. Figure 2.2 shows a further classification of these crosswalks.

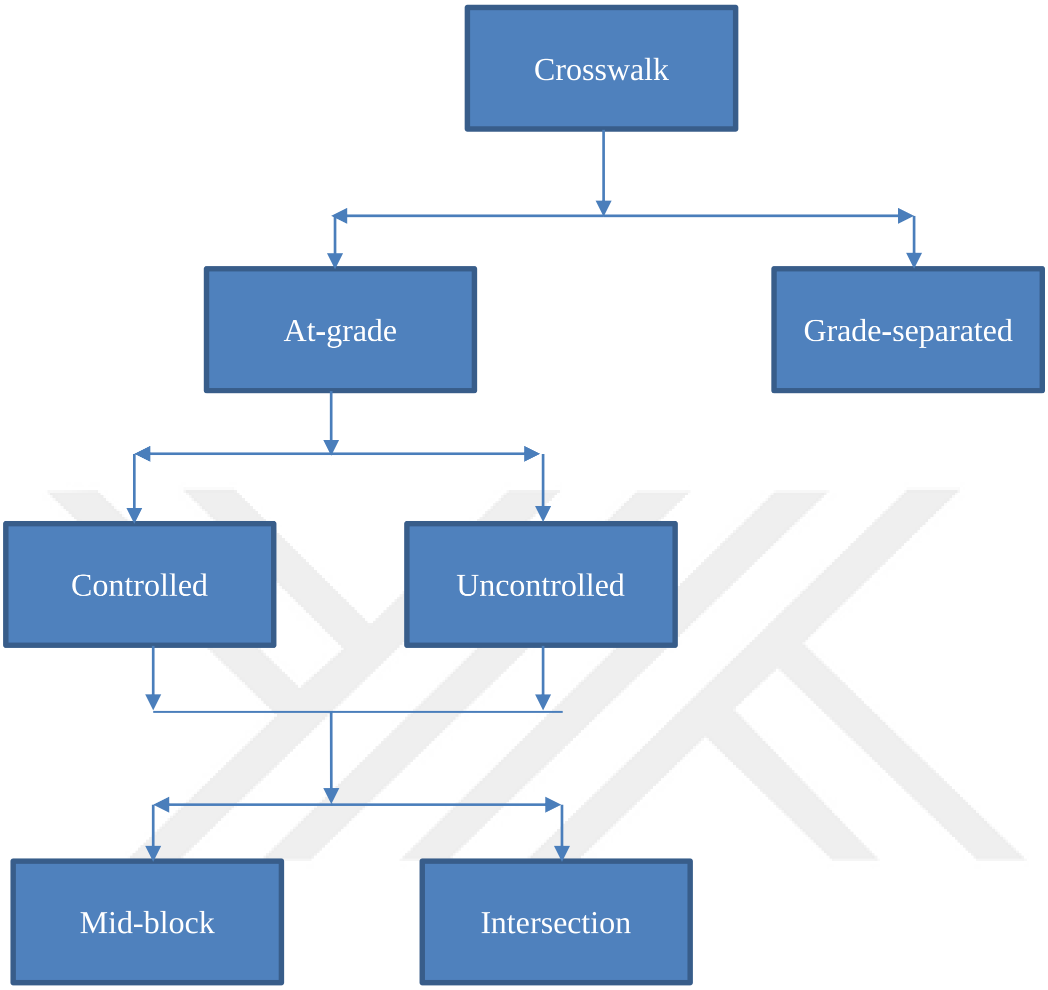

Figure 2.2: The classification of crosswalks (Mitman et al., 2010).

At at-grade crosswalks, the pedestrians crossing the road and vehicles on the road are at the same level (Mitman et al., 2010). Depending on the traffic control elements, there are two types of at-grade crossings: controlled and uncontrolled. In controlled crosswalks, there is the traffic signal to manage the traffic on the street. On the other hand, in the uncontrolled crosswalks, the crosswalk is displayed by pavement markings; however, there are no traffic signals (Mitman et al., 2010). Uncontrolled crosswalks also have two different types, which are mid-block and intersection crossings. Mid-block crosswalks are located between intersections, and they usually have pavement markings. Whereas, crosswalks at intersections may be marked or unmarked (Mitman et al., 2010). At grade-separated crosswalks, the pedestrian crosses the road at the different grade from the vehicles. These crosswalks can be in various forms such as pedestrian underpasses or footpaths across the road (Mitman et al., 2010).

The near-side is the side that the right turning vehicles can conflict with pedestrians and the pedestrians whose origin is the near-side is called near-side pedestrians. The near-side and far-side pedestrians at an intersection crosswalk are shown in Figure 2.3 as point A and B, respectively.

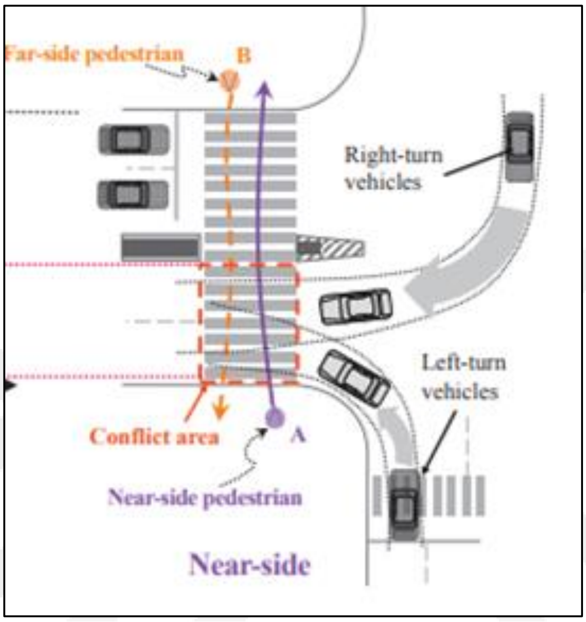

Figure 2.3: Near-side and far-side pedestrians at an intersection crosswalk (Iryo-Asano and Alhajyaseen, 2014).

The near-side pedestrians can pass through the conflict area (where paths of pedestrians and vehicles overlap) in the first half of crossing. Hence, the necessary clearance time for them is the time required to go through half of the crosswalk in this case (Iryo-Asano and Alhajyaseen, 2014). The pedestrians whose origin is the far-side is called far-side pedestrians, and they need to complete crossing to pass the conflict area. Consequently, they need sufficient clearance time to cross the entire crosswalk (Iryo-Asano and Alhajyaseen, 2014).

Regardless of the crosswalk type, the crossing behavior of pedestrians follows a pattern with three distinct levels. These levels are strategic, tactical and operational levels (Ishaque et al., 2008). The strategic level occurs before the pedestrians reach the crosswalk and it is mostly related to the decision to walk. At this level, the pedestrians choose to or not to walk depending on the activity they want to fulfill. This level is not directly related to the crossing action, but it indirectly paves the way for it. At the tactical level, the pedestrian makes short-term decisions about the priority of activities, where the activities are and choosing the right route to reach the activities (Ishaque et al., 2008). There is a two-way relation between the first and second level. Because the decision made at the tactical level is related to the decision made at the strategic level and it has an impact on it (Ishaque et al., 2008). For example, the mode choice is the strategic choice. This choice is related to the tactical level, which designates the crossing speed on the street or even taking the risk of crossing at an uncontrolled or signalized intersection.

At the operational level, the path planning from the present location to the next activity location is performed. The destination at the operational level is determined, and it is affected by the activity at the tactical level. At the operational level, the pedestrian decisions can be studied concerning two aspects. The first one is the behavioral aspect and the second one is the engineering aspect. The behavioral aspect focuses on the pedestrians decisional behavior for crossing, their variation in their speed and the effect of other pedestrians (Ishaque et al., 2008). The engineering aspect considers issues such as the effect of traffic control measures on pedestrian behavior. The immediate decision to cross is made at the operational level. This decision is related to the characteristics of the pedestrian such as the walking speed that can be low or fast. Figure 2.4 presents the pedestrian behavior levels.

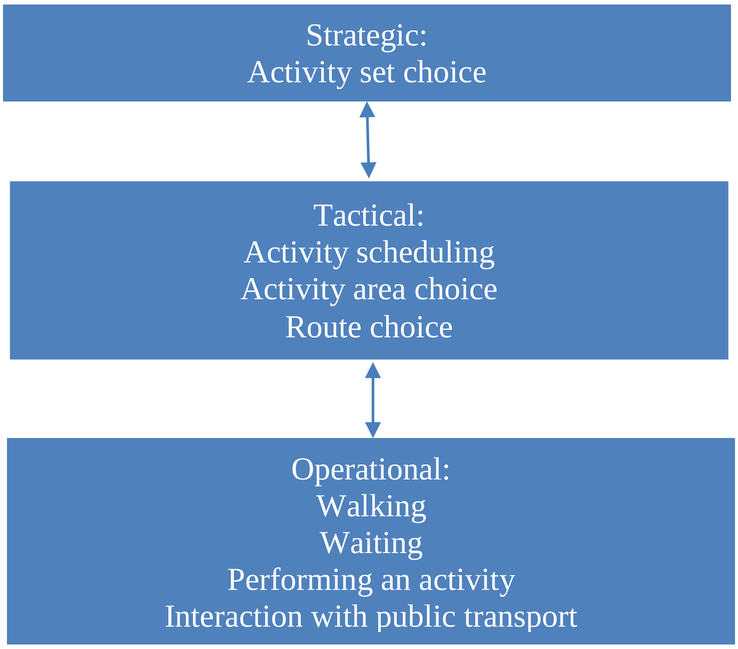

Figure 2.4: Pedestrian behavior levels (Ishaque et al., 2008).

The operational level is affected by the strategic and tactical levels, which the decisions are made. For instance, when time-saving is essential, a pedestrian could decide to walk faster; the decision made at the operational level. In this context, the choice to walk decision is a strategic level decision, and the route choice is a decision at the tactical level (Ishaque et al., 2008).

In the literature, some studies have provided essential information about the effect of pedestrian demographic characteristics such as gender and age on the behavior while crossing the road. These studies generally focused on the effect of pedestrians properties on their road crossing decision by considering the speed or distance of vehicles and other important factors, which are presented here chronologically.

Oxley et al. (1997) focused on the elderly pedestrians to understand that whether they because of declines in their physical and perceptual abilities, are more vulnerable to crashes or not. They used video recording to observe the pedestrians road crossing behavior at several urban locations. The average kerb delays (the time interval between the back of the last vehicle passed the pedestrian and the pedestrian's start to cross) of young and the elderly pedestrians were measured. Also, the gap acceptance (the distance of a near-side oncoming vehicle from a pedestrian at the time of the first step forward to cross the road) of young and the elderly pedestrians on one-way and the two-way roads were found. Table 2.1 shows the result of their study for the elderly and young pedestrians.

Table 2.1: The average kerb delay and gap acceptance of young and the elderly pedestrians on one-way and two-way roads (Oxley et al., 1997).

Age	One-way roads		Two-way roads	
	Average kerb delay (s)	Gap acceptance (m)	Average kerb delay (s)	Gap acceptance (m)
Elderly	-0.10	134.10	-0.87	69.10
Young	-0.11	119.20	-0.05	51.30

The results of this study indicated that on one-way roads, the elderly pedestrians' crossing behavior is similar to the young counterparts and it is considerably safer than their crossing on the two-way roads. Moreover, the results showed that the age-related perceptual deficits are an important factor which increases the risk of involving in an accident.

In a study by Hamed (2001), pedestrian crossing behavior models at mid-block crosswalks on undivided and divided roads at Amman, Jordan were estimated. It was revealed that the pedestrian waiting time has an impact on the number of attempts needed to cross the street successfully. Also, the results show that pedestrians have different behaviors while crossing the road from one side to the middle part and from the middle section to the other side on undivided roads. On divided roads, male pedestrians are 1.35 times more likely to have a less waiting time for crossing from one side to the refuge than females. Also, males are 3.105 times more likely to have less waiting time than females for crossing from the refuge to the other side of the street.

In the study by Oxley et al. (2005), simulated road crossing task was used to study the effects of some factors on pedestrians crossing decision that was related to time gap (the time interval between the pedestrians who want to start to cross the street and the upstream vehicle arriving in the first lane and has not been crossed in front of the pedestrian). Pedestrians crossing decisions were analyzed and the impact of the age, time gap, speed of vehicles, distance of the oncoming vehicle and walking time is studied. Results showed that pedestrian crossing decision was mostly affected by the distance of the oncoming vehicle. Furthermore, elderly pedestrians did not choose the proper time gap compared to young pedestrians. These results show the importance of the time gap in the pedestrian crossing choice behavior and the effect of age on choosing the time gap. The proportion of yes responses (accept) to cross the street as a function of age group, distance gap, vehicle speed, and time gap is shown in Figure 2.5, Figure 2.6 and Figure 2.7 for three different age groups including young, young-old and old-old groups, respectively.

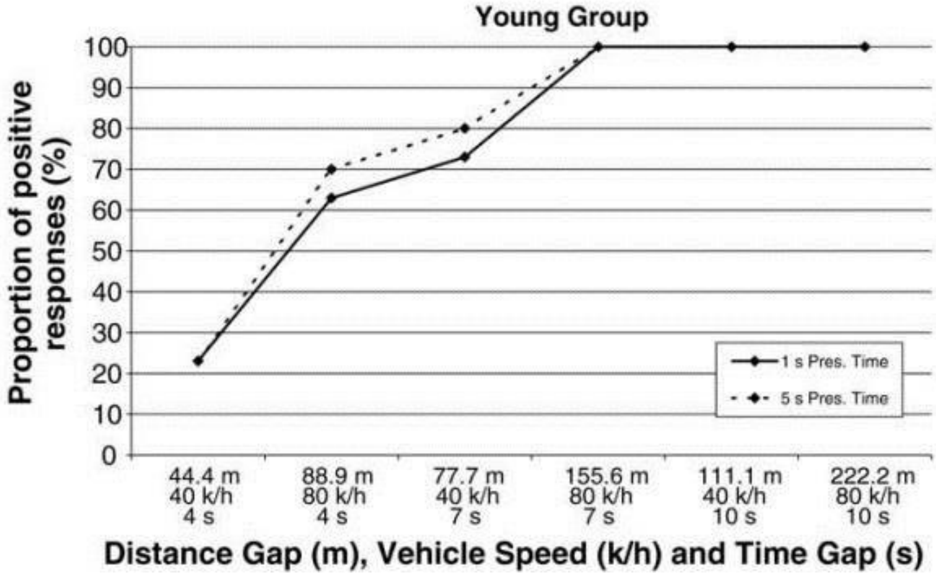

Figure 2.5: The proportion of yes responses of the young group in different situations (Oxley et al., 2005).

The responses of the elderly participants were more affected by both the time gap and distance factors. For instance, when the distance is increased, the response rate rose from 44.4 m to 88.9 m and from 77 m to 155.6 m, and vehicle speed increased from 40 km/h to 80 km/h, despite time gaps being constant at 4 s and 7 s, respectively. However, response rates also increased when time gaps increased from 4 s to 7 s and stayed at asymptote.

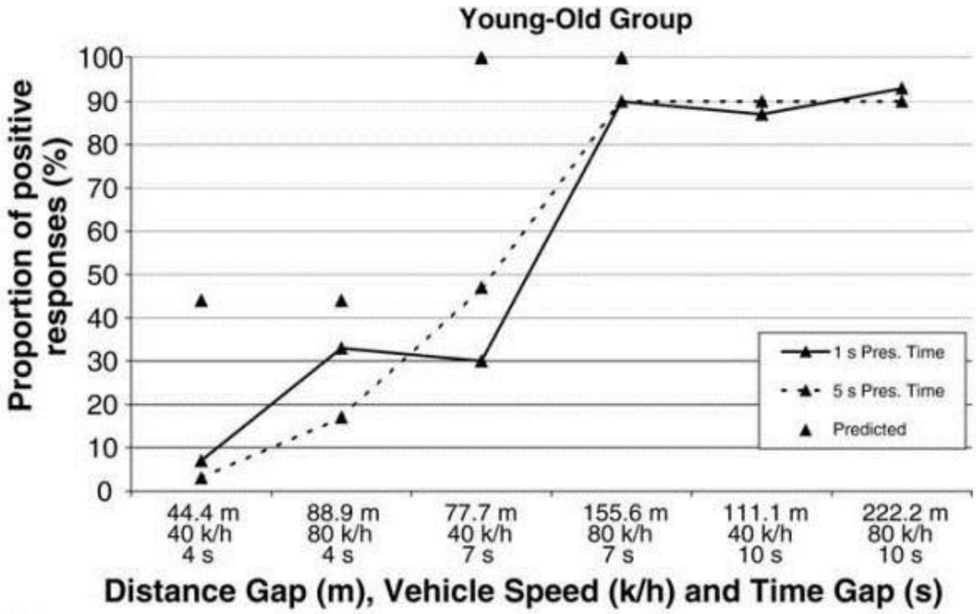

Figure 2.6: The proportion of yes responses of the young-old group in different situations (Oxley et al., 2005).

They found that even when old–old adults had adequate time to process the time gap of oncoming vehicles, but many of them based on vehicle distance, made unsafe crossing decisions.

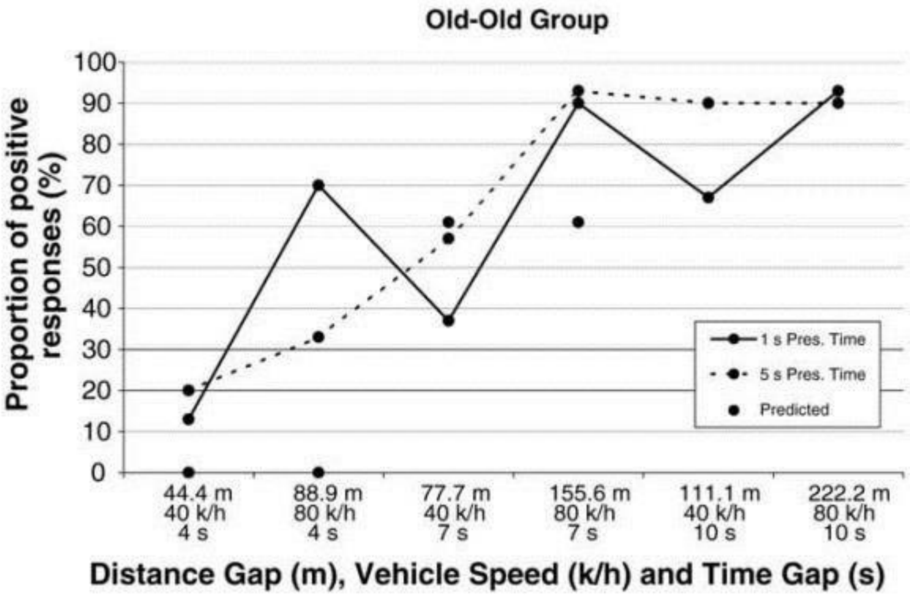

Figure 2.7: The proportion of yes responses of the old-old group in different situations (Oxley et al., 2005).

In a study by Ishaque and Noland (2008), pedestrians crossing behavior and their choice of speed at the micro-scale were focused. It was shown that pedestrian choice of speed was related to the risk of the specific situation, individual capabilities, and value of time. By making pavement surfaces smoother and safer for elderly pedestrians, their capability can be improved. The pedestrian capability interacts with the risk associated with traffic. When the risk of the accident reduces, the delay due to the risk avoidance will decrease which leads to having less travel time. By increasing pedestrian signal cycle times and slowing traffic, the older people tend to wait for longer gaps in traffic. The pedestrian who has a higher value of time, have higher speeds, especially during peak hours. Moreover, the risk-taker pedestrians have higher crossing speed than other pedestrians. Also, the decision-making process was formalized by specifying the choice of speed (S) based on the capabilities of the pedestrian (C), their value of time (V), and the potential risk (R) as shown in equation (2.1).

$$S = f\,(C, V, R) \tag{2.1}$$

Bernhoft and Carstensen (2008) focused on comparing the behavior and preferences of the elderly pedestrians with the group of people aged 40–49 in cities. It was found that elderly pedestrians appreciate pedestrian crossings and signalized intersections more than the younger pedestrians. Figure 2.8 shows the behavior of young and elderly pedestrians in various traffic situations.

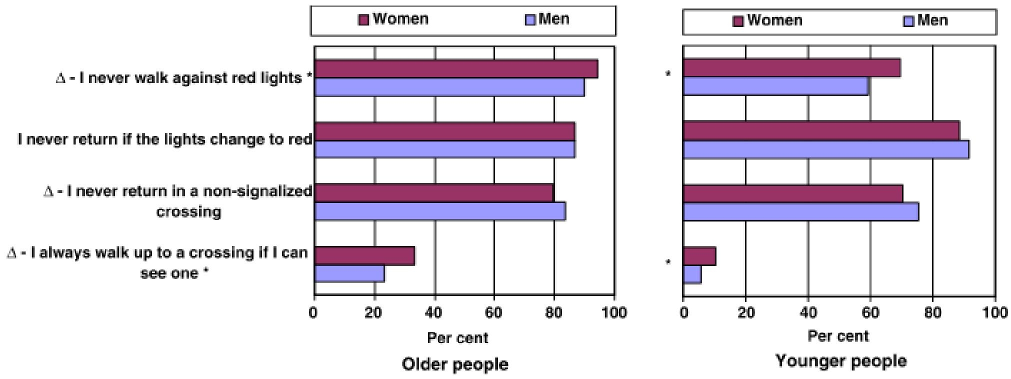

Figure 2.8: Young and elderly pedestrians' behavior in various traffic situations (Bernhoft and Carstensen, 2008).

When the road is without the facilities, the elderly pedestrians feel that it is dangerous to cross the road. This behavior difference can be related to differences in physical abilities and health rather than to differences in age and gender.

The proportion of the elderly pedestrians is more than the proportion of young pedestrians for cases of always deciding to walk up to a pedestrian crossing if they can see one, never crossing at a red light and never returning in a non-signalized crossing as presented in Figure 2.8. Hence, elderly pedestrians have more cautious behavior.

In a study by Jamil et al. (2015), the pedestrian crossing choice models according to road, traffic and human factors were studied. The results showed that the pedestrian crossing choice was significantly affected by traffic flow and road type. Also, it was found that human factors have more effect than the mentioned factors. Three kinds of pedestrians were introduced which were named risk-taking pedestrians, conservative pedestrians, and pedestrian for pleasure.

Jamil et al. (2015) worked on the uncontrolled marked crosswalk when the pedestrians crossed the road with changing directions and speeds that result in curved paths and higher chances of safety issues. They analyzed the pedestrians crossing patterns concerning entry/exit pairs, and also the turning points. The results showed that avoiding collision with grouping pedestrian crossing together from the same or opposite direction, short or also fast distance and avoiding traffic running straight on the lanes have the highest effect on the pedestrian crossing and these results in curved paths. The results are shown in Table 2.2.

Table 2.2: The percentage of each reason for curved crossing patterns (Jamil et al., 2015).

Factors	Percentage
Forced by the traffic flow of the same direction	7.31
Avoiding collision with grouping pedestrian crossing together from the opposite direction	13.85
Avoid collision with pedestrians waiting at the edge of the zebra	0.77
Avoiding collision with the right turning vehicles	0.77
Avoiding traffic running straight on the lanes	0.77
Short distance	71.15
Pedestrian following plus avoiding stopped vehicles illegally on the zebra	5.38

Ferenchak (2016) focused on the relation between pedestrians behavior with motor vehicles. The data about pedestrians characteristics were gathered from a mid-block crossing in Bangalore, Karnataka, India. By using logistic regression, it was found that the pedestrian's waiting time increases as the pedestrian gets elder. Moreover, elderly pedestrians have fewer conflicts with motor vehicles compared to younger pedestrians while crossing the road. Furthermore, it is found that males caused conflicts with motor vehicles two times more than females. Also, the waiting time of males was about half of the waiting time of females. Table 2.3 presents the differences in pedestrian behavior by gender with the statistical significance of the differences.

Table 2.3: Pedestrians' crossing behavior according to their gender (Ferenchak, 2016).

	Male	Female	The difference (%)	P-value
Waiting time (s)	18.33	34.40	88.00	0.00
Utilization of crosswalk (%)	39.80	53.20	33.70	0.06
Causing a conflict (%)	33.10	14.3	131.5	0.00

Table 2.4 shows the result of the logistic regression between the gender/age and risk/conflict.

Table 2.4: Logistic regression between gender/age and risk/conflict (male = 0) (Ferenchak, 2016).

Factors	Coefficient	Standard error	P-value
Gender			
Utilization of crosswalk	0.543	0.298	0.067
Causing a conflict	-1.086	0.380	0.004
Age			
Utilization of crosswalk	0.034	0.012	0.005
Causing a conflict	-0.046	0.015	0.002

As indicated in Table 2.3, the female and male samples were statistically different from one another concerning the waiting time. Also, males were less likely to use the crossing infrastructure than females properly. However, as seen in Table 2.4, this relationship was not significant. Moreover, the probability of causing a conflict reduces as age increases. This relationship is statistically significant at a 99% confidence level as well.

3. DATA

In this thesis, the data are collected from an uncontrolled crosswalk on Aytar Street in Istanbul-Turkey shown in Figure 3.1 and Figure 3.2. The data are gathered by recording pedestrians on the crosswalk for 12 hours on a typical working day. The Aytar Street is a one-way street which has two lanes, and the width of the street is 8 m. The parking is prohibited on both sides of the street. However, there are sometimes vehicles parked illegally.

Figure 3.1: Aytar Street in Istanbul-Turkey (Google Map, 2018).

Figure 3.2: Crosswalk on Aytar Street in Istanbul-Turkey.

The characteristics of the pedestrians and traffic at the uncontrolled crosswalk are observed to analyze the pedestrians crossing behavior. By using the camera records, the pedestrians' characteristics including gender, age, using the mobile phone while waiting for crossing the street, the number of pedestrians who start waiting at the same time, pedestrians with kids and/or carrying something are observed. Here, age is distinguished for being elderly or not through appearance.

The second group of observations is about pedestrian/traffic interaction. For this group, the gap at the crossing, waiting time and average headway of rejected vehicles are observed. The gap is measured at a section on the crosswalk. It is the time between the last vehicle before (back bumper), and the first vehicle (front bumper) after the pedestrian crosses the street. Also, the pedestrians waiting time is measured which is started when the pedestrian approaches the pavement until the pedestrian sets foot on the street to cross it. The average headway of rejected vehicles by the pedestrian to cross is found which is between 1.25s to 5.01s and when the pedestrians waiting time is 0, the average headway of rejected vehicles is assumed to be equal to 0 in the dataset. The term "headway" is defined as the time between the front bumper of one vehicle and the front bumper of the next (Hutchinson, 2008). The number of rejected vehicles by pedestrians for crossing the street is observed which is between 1 to 16 vehicles and for the pedestrians with zero waiting time, the number of rejected vehicles for them is assumed to be equal to 0 in the dataset. Figure 3.3 clearly shows the difference between the gap and the headway.

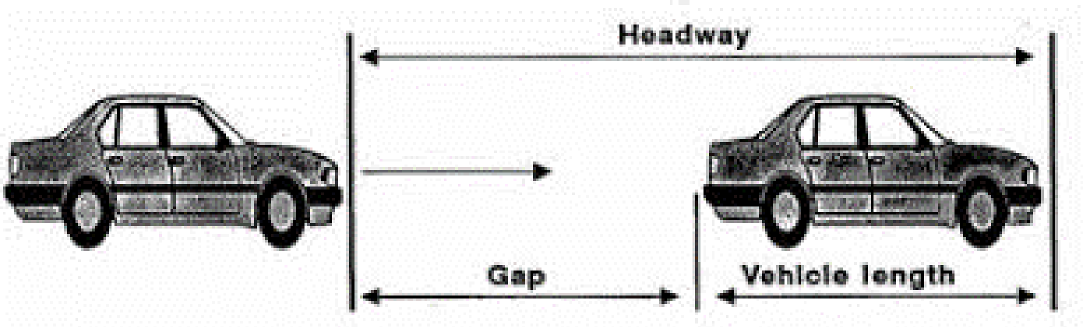

Figure 3.3: The difference between gap and headway (Fadilah and Shariff, 2014).

The final set of observations is about the traffic attributes. The observations include illegal parking at the crosswalk, the cases if the driver approaching yield to stop or reduce speed or change course.

Table 3.1 presents the averages of the gap, waiting time, average headway of rejected vehicles, the number of rejected vehicles and the number of pedestrians start waiting at the same time (PWS).

Table 3.1: The average of each Factor.

Factor	Average
PWS	1.39
Gap (s)	6.22
Waiting Time (s)	6.41
Average headway of rejected vehicles (s)	1.45
Number of rejected vehicles	1.47

The average headway of rejected vehicles in Table 3.1 is observed to take values as high as 5.01 s while maximum PWS is four. According to the Table, the pedestrians at the observed crosswalk reject at least one vehicle (wait for at least one vehicle) before crossing. Moreover, the crosswalk and traffic conditions seem to make pedestrians wait relatively longer since average PWS is about one.

Table 3.2 shows the number and percentage of pedestrians for each group of categories including gender, age, using the phone, approaching driver yielding, PWS and crossing at the first time.

Table 3.2: Number and percentage of pedestrians in each pedestrians group.

Factor	Group	#	Percentage
Gender	Male	359	58.09
	Female	259	41.91
Age	Young	584	94.5
	Elderly	34	5.50
Using phone	Using	38	6.15
	Not using	580	93.85
Driver yielding	Yes	276	44.66
	No	342	55.34
Crossing at the first attempt	Yes	354	57
	No	264	43
PWS	Alone	434	70.23
	2	136	22.01
	3	36	5.82
	4	12	1.94

Table 3.3 and 3.4 presents the average pedestrians' gap, waiting time and average headway of rejected vehicles for each pedestrians group and pedestrians characteristics, respectively.

Table 3.3: The average pedestrians' gap, waiting time and average headway of rejected vehicles for each pedestrian group.

Factor	Group	Gap (s)	Waiting time (s)	Average headway of rejected vehicles (s)
Gender	Male	6.16	5.57	1.26
	Female	6.31	7.57	1.71
Age	Young	6.19	6.18	1.41
	Elderly	6.78	10.42	2.11
Using phone	Using	6.41	8.36	1.48
	Not using	6.21	6.28	1.45
Driver yielding	Yes	6.38	6.60	1.49
	No	6.08	6.26	1.42
PWS	Alone	1.35	6.93	1.50
	2	1.16	5.71	1.45
	3	1.10	2.55	0.99
	4	1.27	7.21	1.08

As it is seen in Table 3.2, males wait shorter and have less average headway of rejected vehicles compared to the females while they also reject at least one vehicle on average. Moreover, elderly pedestrians' waiting time and average headway of rejected vehicles are higher than the young pedestrians. Also, the pedestrians who use the mobile phone while waiting for crossing the street have higher waiting time and average headway of rejected vehicles than who do not use it. Furthermore, when driver approaching to stop or reduce speed, the pedestrians' waiting time and average headway of rejected vehicles are higher. The PWS range in the dataset is between 1 (pedestrian start waiting alone) and 4 and the minimum and maximum waiting time is when PWS is 3 and 4, respectively. However, the minimum and maximum average headway of rejected vehicles are when PWS is 3 and 1 (pedestrian start waiting alone), respectively.

Table 3.4: The average pedestrians' gap, waiting time and average headway of rejected vehicles for each pedestrian characteristics.

Pedestrian category	Group	Gap (s)	Waiting time (s)	Average headway of rejected vehicles (s)
Young	Male	6.10	5.08	1.19
	Female	6.32	7.66	1.71
Elder	Male	7.03	12.47	2.31
	Female	6.17	5.50	1.63

Table 3.4 shows that the elderly male's average headway of rejected vehicles and waiting time is more than the elderly females. Also, the young males have shorter waiting time and average headway of rejected vehicles than the young females.

Figure 3.4 shows the histogram of waiting time to understand the frequency of waiting time in the interval of 5s.

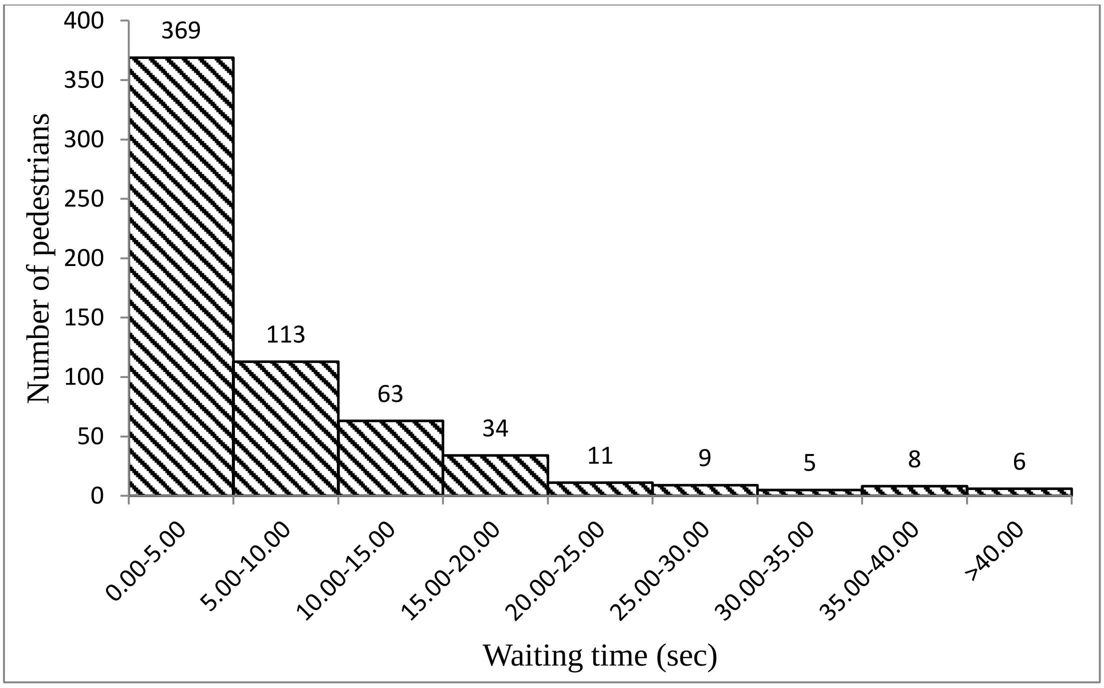

Figure 3.4: Histogram of waiting time.

Figure 3.4 clearly shows that the majority of pedestrians' waiting time is between 0 s and 5 s.

3.1 Statistical distribution of data

In this section, the aim is understanding whether the distribution of the waiting time of each group of factors is normally distributed or not. For this investigation, the chi-square test is used. The chi-square test statistic (${\chi_{stat}}^2$) is calculated by equation 3.1 (Bury, 1999).

$$ {\chi_{stat}}^2 = \sum_{i=1}^{n} \frac{(E_i - O_i)^2}{(E_i)} \qquad (3.1) $$

Where:
n: Number of classes
E_i: The expected frequency of type i

O_i: The observed frequency.

The class ranges are determined with equal probability. The number of classes is calculated by the number of data, which can be followed by equation 3.2 (Bury, 1999).

$$ m = 1 + 3.3 * \log(n) \qquad (3.2) $$

Where m is the number of classes and n is the number of pedestrians. The degree of freedom (df) is calculated by equation (3.3).

$$ df = m - n_p - 1 \qquad (3.3) $$

Where:
m: Number of intervals
n_p: Number of parameters of the selected distribution
$m \geq 5$ & $n \geq 5$

The number of parameters of the normal distribution is 2 (Bury, 1999). By calculating the df and using 5% level of significance, the critical value of chi-square is found from the chi-square distribution table. The calculated sum of squared deviations of observed and theoretical frequencies for normal distribution model is presented in Table 3.5. When the chi-square statistic is bigger than the critical value, it means that the distribution is not fitted to the normal distribution; otherwise, it is fitted to the normal distribution.

Table 3.5: Distribution models and their probability density functions (Bury, 1999).

Distribution model	Probability density function	Parameters	Mean	Variance
Normal	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma 2}}$	μ, σ	μ	σ^2

Table 3.6 shows the results of the χ^2 test to investigate whether the distribution of the waiting time for each group of each factor is the normal distribution or not.

Table 3.6 shows that the distribution of the waiting time of pedestrians who carry something while crossing the street and who using the mobile phone while waiting for crossing the street is the normal distribution as the chi-square statistical value ($\chi^2{}_{statistical}$) is less than the chi-square critical value. However, the distribution of other groups of factors is not normally distributed, as the $\chi^2{}_{statistical}$ is bigger than the chi-square critical value.

Table 3.6: The results of the χ^2 test for distribution of waiting time of each group of each factor.

Factors	Group	df	$\chi^2{}_{critical}$	$\chi^2{}_{statistical}$	P-Value	Distribution of waiting time
Driver yielding	Yes	8	15.51	50.91	2.7307E-8	Not normal
	No	8	15.507	34.929	2.7542E-5	Not normal
Crossing diagonally	Yes	4	9.4877	9.9813	0.04074	Not normal
	No	9	16.919	75.284	1.3887E-12	Not normal
Carrying something	Yes	1	3.8415	0.64627	0.42145	Normal
	No	9	16.919	86.898	6.7724E-15	Not normal
Gender	Male	8	16.919	89.045	1.0170E-4	Not normal
	Female	9	15.507	31.787	2.5535E-15	Not normal
Age	Elderly	2	5.9915	8.6824	0.01302	Not normal
	Young	9	16.919	89.045	2.5535E-15	Not normal
Using phone	Using	2	5.9915	4.8045	0.09051	Normal
	Not using	9	16.919	89.045	2.5535E-15	Not normal
Presence of illegally parked vehicle	Yes	9	16.919	89.045	2.5535E-15	Not normal
	No	8	15.507	60.353	3.9739E-10	Not normal

3.2 Comparing Population Mean

In order to assess the statistical differences between groups of each factor, there are some tests such as pooled t-statistic, non-pooled t-test (Welch's t-test) and Mann-Whitney test. The t-statistic is used when the distribution of the response variable of both groups is the normal distribution with equal variance (Homoscedasticity). The non-pooled t-test (Welch's t-test) is considered when the distribution of the response variable of both groups is the normal distribution with unequal variance (Heteroscedasticity). Mann-Whitney test is used when the waiting time distribution of at least one of the groups is not normally distributed with the same shape. When the sign of the skewness and kurtosis are the same for both groups of the factor, their shape is same (Bury, 1999).

To measure the degree of asymmetry of a frequency distribution the skewness is calculated by equation (3.4) (Bury, 1999).

$$Skewness = \frac{n}{(n-1) \times (n-2)} \times \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{s^3} \tag{3.4}$$

Where,

n: Number of observations

s: Sample standard deviation

$\bar{x}$: Sample mean

The sample standard deviation shows how closely the values of a data set are clustered around the mean, and it is obtained by equation (3.5) (Bury, 1999).

$$\text{Sample standard deviation} = s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} \tag{3.5}$$

Where,

n: Number of observations

$\bar{x}$: Sample mean

In order to measure the flatness of a frequency distribution, the kurtosis is calculated which is obtained by equation (3.6) (Bury, 1999).

$$\text{Kurtosis} = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{s^4}\right] - \frac{3\,(n-1)^2}{(n-2)(n-3)} \quad (3.6)$$

Table 3.7 presents the statistic of pedestrians waiting time in each group of each factor and the proper method for each factor (Burry, 1999).

Table 3.7: Statistic of waiting time of each group and suitable method for each factor.

Factors	Group	Skew-ness	Kurtosis	Same shape	Type of distribution	Proper test
Driver yielding	Yes	2.65	3.03	Yes	Not normal	Mann-Whitney U
	No	9.07	14.30		Not normal	
Crossing diagonally	Yes	2.07	4.73	Yes	Normal	Mann-Whitney U
	No	2.83	11.17		Not normal	
Carrying something	Yes	1.53	2.02	Yes	Not normal	Mann-Whitney U
	No	2.84	11.32		Not normal	
Gender	Male	2.85	11.43	Yes	Not normal	Mann-Whitney U
	Female	2.17	6.23		Not normal	
Age	Elderly	2.85	2.62	Yes	Not normal	Mann-Whitney U
	Young	11.43	9.86		Not normal	
Using phone	Using	1.58	1.62	Yes	Normal	Mann-Whitney U
	Not using	2.85	11.43		Not normal	
Presence of illegally parked vehicle	Yes	11.43	3.07	Yes	Not normal	Mann-Whitney U
	No	2.85	3.07		Not normal	

As it is indicated in Table 3.7, the Mann-Whitney U test can be used for all factors.

3.2.1 Comparision of two populations with Mann-Whitney U test

The Mann-Whitney U test is a nonparametric test of the null hypothesis that two samples come from the same population against an alternative hypothesis that two samples are from different populations. Unlike the t-test, it does not require the assumption of normal distributions. In the Mann-Whitney U test, when the sample sizes are big ($n > 20$), the value of U approaches to normal distribution, and so the null hypothesis should be tested by Z-test or P-value. In order to calculate the U statistic (U_{sta}), first, all the data must be ranked together; ignoring which group they belong to it. Then each group ranks must be added up (T_1, T_2) and U_{sta} will be calculated with equation 3.7 (Bury, 1999):

$$U_{sta} = \begin{cases} U_1 = (n_1 * n_2) + \left(n_1 * \frac{n_1 + 1}{2}\right) - T_1 \\ U_1 = (n_1 * n_2) + \left(n_2 * \frac{n_2 + 1}{2}\right) - T_2 \end{cases} \tag{3.7}$$

Where:

n_1 and n_2: Number of data in the first and second group, respectively.

T_1 and T_2: Sum of the ranks in the first and second group, respectively.

After calculating the U_{sta}, the value of Z_{sta} can be calculated with equation 3.8 (Bury, 1999) and to test the null hypothesis, it must be compared to the critical Z value ($Z_{critical}$) obtained from Z table for the assumed level of significance which is 0.05 for this study.

$$U_{sta} \rightarrow Z_{sta} \Rightarrow Z_{sta} = \frac{U_{sta} - \mu_u}{\sigma_u} \tag{3.8}$$

$$\sigma_u = \sqrt{\frac{n_1 n_2 * (n_1 + n_2 + 1)}{12}} \quad , \quad \mu_u = \frac{n_1 * n_2}{2}$$

The two different groups for each characteristic do not come from the same population if their Z_{sta} does not include in the range of the $Z_{critical}$; otherwise, they come from the same population and cannot be separated as two different groups. The results of Mann-Whitney U test at 5% level of significance are shown in Table 3.8.

Table 3.8: Results of Mann-Whitney U test for waiting time of pedestrian.

Characteristic	Groups	Z_{sta}	$Z_{critical}$	Decision
Gender	Male Female	-3.84	±1.96	Not same population
Using phone	Using Not using	-0.33	±1.96	Same population
Age	Elderly young	-2.18	±1.96	Not the same population
Presence of illegally parked vehicle	Yes No	-0.25	±1.96	Same population
Crossing diagonally	Yes No	-1.25	±1.96	Same population
Carrying something	Yes No	-1.28	±1.96	Same population
Driver yielding	Yes No	-0.72	±1.96	Same population

The results given in Table 3.8 indicate that males and females for gender characteristic do not come from the same population because Z_{sta} does not include in the range of the $Z_{critical}$. Also, pedestrians who are the elderly pedestrian or young, do not come from the same population. The other groups in each characteristic come from the same population because their Z_{sta} included in the range of the $Z_{critical}$ which shows that it cannot be considered as two separate groups. Hence, gender can be considered as two groups including male and female. Moreover, age can be classified into two groups which are the elderly pedestrian and young pedestrian.

3.2.2 Comparison of multiple populations with Kruskal-Wallis H test

Kruskal-Wallis H test is a rank-based non-parametric hypothesis test, which is not assumed a normal distribution of the dataset. The null hypothesis in this test is "all groups are from the same population," and the alternative hypothesis is "at least one group is from the different population." It is used to indicate that whether there are statistically significant differences between two or more groups of independent variables or not. This test is considered the non-parametric alternative to the one-way ANOVA test to allow the comparison of more than two independent groups which do not have the normal distribution, but they have equal variance. If in each factor, the standard deviation of one group is equal or less than two times the standard deviation of the other groups, the variance of these groups is equal.

Table 3.9 shows whether the Kruskal-Wallis H test is proper for each factor or not.

Table 3.9: Statistic of waiting time of each group and finding the new proper test for each factor.

Factors	Group	Standard deviation	Equal Variance	Type of distribution	Proper test
Driver yielding	Yes	9.82	Yes	Not normal	Kruskal-Wallis H
	No	8.33		Not normal	
Crossing diagonally	Yes	6.42	Yes	Normal	Kruskal-Wallis H
	No	9.23		Not normal	
Carrying something	Yes	5.71	Yes	Not normal	Kruskal-Wallis H
	No	9.07		Not normal	
Gender	Male	8.83	Yes	Not normal	Kruskal-Wallis H
	Female	9.17		Not normal	
Age	Elderly	15.76	Yes	Not normal	Kruskal-Wallis H
	Young	8.53		Not normal	
Using phone	Using	11.22	Yes	Normal	Kruskal-Wallis H
	Not using	8.85		Not normal	
Presence of illegally parked vehicle	Yes	8.52	Yes	Not normal	Kruskal-Wallis H
	No	9.17		Not normal	

As it is mentioned in Table 3.9, all factors have at least one group with not the normal distribution, but with equal variance. The groups in each factor have equal variance, because in each factor, the standard deviation of one of the groups is not more than two times of the standard deviations of other groups, or two times of the standard deviation of one of the groups is not less than the standard deviations of other groups. Hence, the Kruskal-Wallis H test can be used for all factors.

When there are more than five data in each group, Kruskal-Wallis H test statistic approximately behaves like chi-square distribution, having df equal to the number of groups which in this study, is two minus one that is equal to one. By knowing the df which is equal to one and assuming the level of significance equal to 0.05, the critical value of the Kruskal-Wallis H test is obtained from the chi-square table. At the first step, all the data must be ranked; ignoring which group, they belong to it. Then the total ranks for each group must be calculated (T_1, T_2,…) and H statistic (H_{sta}) will be calculated using equation (3.9) (Bury, 1999):

$$H_{sta} = \left[\frac{12}{N*(N+1)} * \sum_{i=1}^{m} \frac{{T_i}^2}{n_i}\right] - 3*(N+1) \tag{3.9}$$

Where:

N: The total number of data in all groups

m: Number of groups

T_i: Sum of ranks in group i

n_i: Number of data in group i

Table 3.10 presents the results of the Kruskal-Wallis H test.

Table 3.10: Results of the Kruskal-Wallis H test for waiting time.

Characteristic	Groups	H(statistical)	H(critical)	Decision
Gender	Male Female	14.724	3.841	Not same population
Using phone	Using Not using	0.111	3.841	Same population
Age	Elderly Young	4.732	3.841	Not the same population
Presence of illegally parked vehicle	Yes No	0.062	3.841	Same population
Crossing diagonally	Yes No	1.555	3.841	Same population
Carrying something	Yes No	1.647	3.841	Same population
Driver yielding	Yes No	0.522	3.841	Same population

Calculated values of Kruskal-Wallis H test given in Table 3.10 indicate that female and male for gender characteristic do not come from the same population because the

H (statistical) is more than the H (critical). Moreover, elderly pedestrians and young pedestrians do not come from the same population. The other groups in each characteristic come from the same population because the H (critical) is higher than the H (statistical) which shows that it cannot be considered as two separate groups.

Therefore, the both Mann-Whitney U test and Kruskal-Wallis H test indicate that the pedestrians regarding their gender can be separated to the two groups, which are male and female. Besides, being the elderly factor can be classified into two groups including the elderly and young groups.

4. MODEL

4.1 An Overview of Discrete Choice Model:

Discrete choice model is a behavioral model to show and predict the individuals' behavior. It has been used a lot in econometrics and transportation science (Antonini et al., 2006).

4.1.1 Utility-based choice theory

The utility allows us to rank a series of alternatives and identify the single alternative that will be chosen. As shown in equation (4.1), the utility function U which describes an individual's utility valuation for each alternative has the property that an alternative is chosen if its utility is higher than the utility of all other alternatives in the individual's consideration set, C (Koppelman and Bhat, 2006).

$$\text{If} \qquad U_i(X_i, S_i) \geq U_j(X_j, S_j) \Rightarrow i > j \quad \forall_j \in C \tag{4.1}$$

$i > j$: Alternative i is selected over alternative j

X_i: Characteristics of the alternative i

S_i: Socio-economic variables of the individual i

X_j: Characteristics of the alternative j

S_j: Socio-economic variables of the individual j

The actual utility (U) includes two components includes the deterministic component (v) and error component (ε) as shown in the equation (4.2) (Koppelman and Bhat, 2006). The deterministic component is the utility observed by the analyst. Error component is unknown utility considered by the individual such as feelings habits etc.

$$U = v + \varepsilon \tag{4.2}$$

Attributes of alternatives and characteristics of individuals are included in the utility function. As it stated in equation (4.3), the utility of an alternative i (v_i) is equal to (Koppelman and Bhat, 2006);

$$v_i = v_i(X, S) = \sum a \times X + b \times S \tag{4.3}$$

a and b: Weight factors or coefficient of attributes or taste factors

The deterministic component is also called alternative-specific constants. The constant in the utility is the average amount of factors not included in the deterministic component that contributes to the utility differences between each utility and reference utility. Reference utility is the utility that can only contain the variables with alternative specific coefficients. The coefficient of the variable is called in this name if it appears only in some of the utility functions or has a different coefficient for each utility function. If the coefficient of the independent variable were the same in all utility functions of the model, it would be called the generic variable; otherwise, it would be alternative-specific variables. Also, if the coefficient of the constant is different in the utilities, it is called alternative-specific constants (Koppelman and Bhat, 2006). The deterministic component represents the observed and measured variables (Ben-Akiva and Lerman, 1985).

Dummy variables are qualitative variables. Sometimes quantitative variables can be represented as dummies. There are two types of coding for dummy variables: Dummy coding and effects coding. In dummy coding, the variable can take values either one or zero that each of them has its definition, which is defined by the analyst. Effect coding uses ones, zeros and minus ones to convey all of the necessary information on group membership (Hensher et al., 2005). The dummy coding is slightly more convenient to set up than effects coding (Daly et al., 2016).

4.1.2 The binomial logit model

The most frequently used model for probabilistic choice between two alternatives is the binomial logit model. In this model, the probability that the alternative one is chosen when the choice set consists of two alternatives is calculated by the following equation (4.4) (Horowitz et al., 1986).

$$\mathrm{Pr}_1 = \frac{e^{V_1}}{e^{V_1}+e^{V_2}} \tag{4.4}$$

Where,

Pr_1: The probability of choosing alternative one by the individual.

e: Exponential function.

V_1: The deterministic component of the utility of alternative one

V_2: The deterministic component of the utility of alternative two

The binomial logit model cannot treat the choice between more than two alternatives. In the binomial logit model, the probabilities of choosing alternatives one and two are equal when the deterministic components of the two alternatives' utilities are equal.

4.1.3 The multinomial logit model

The binomial logit model can be extended to accommodate choices between more than two alternatives. The deterministic components of the utilities of the alternatives are V_1, $V_2 \ldots V_J$. Then the probability of choosing alternative i (i=1…, J) is shown in equation (4.5) (Horowitz et al., 1986).

$$\mathrm{Pr}_i = \frac{e^{V_i}}{\sum_{j=1}^{J} e^{V_j}} \tag{4.5}$$

4.1.4 Calculation of the coefficients

The procedure to estimate the maximum likelihood includes two steps which are as follows,

1) Developing a joint probability density function of the observed sample, called the likelihood function,

2) Parameter values estimation which maximizes the likelihood function. The likelihood function for a sample of ‘T' individuals, each with ‘J' alternatives is presented in equation (4.6) (Koppelman and Bhat, 2006).

$$L(\beta) = \prod_{\forall t \in T} \prod_{\forall j \in J} (P_{jt}(\beta))^{\delta_{jt}} \tag{4.6}$$

Where,

δ_{jt}= 1 is chosen indicator; it is equal to 1 if j is chosen by individual t and otherwise it is equal to 0. The P_{jt} is the probability that alternative j is chosen by individual t.

By calculating the first derivative of the likelihood function and equating it to zero, the values of the parameters which maximize the likelihood function are found. The log-likelihood function is maximized instead of the likelihood function itself. Because the log of a function is easier to differentiate and it yields the same maximum as the function. The log-likelihood function is expressed in equation (4.7) (Koppelman and Bhat, 2006).

$$LL(\beta) = \mathrm{Log}(L(\beta)) = \sum_{\forall t \in T} \sum_{\forall j \in J} \delta_{jt} \times \ln(P_{jt}(\beta)) \quad (4.7)$$

4.1.5 Comparison method

To evaluate whether the estimated model is an improved model compared to a base model or not, the likelihood ratio test can be used as shown in equation (4.8). The utility of the base model includes only constants. The utility function of the base model has two types which are market share model (LL(M)) and null model LL(0). The market share model consists only constants without any independent variables. The other model is the null model in which all parameters are set equal to zero. The log-likelihood of the base model is less than the log-likelihood of the estimated model.

$$-2\mathrm{LL} = -2 \times (\mathrm{LL}_{\mathrm{Base}} - \mathrm{LL}_{\mathrm{estimated}}) \quad (4.8)$$

Where,

-2LL: Log-likelihood ratio

LL: Log-Likelihood

The calculated test statistic is chi-square distributed with degrees of freedom equal to the difference between the numbers of parameters estimated between the two models. After calculating the df, the critical chi-square value can be determined from the chi-square table depends on the level of the significance. If the -2LL value is larger than the critical chi-square value, the model, which has the grater LL, is improved compared to the model, which has the smaller LL. If the -2LL value is less than the critical chi-square value, the model, which has the grater LL, is not improved compared to the model, which has the less LL.

The likelihood ratio test can also be used for comparing two estimated models by a similar equation (4.9) (Koppelman and Bhat, 2006).

$$-2\text{LL} = -2 \times (\text{LL}_{\text{Small}} - \text{LL}_{\text{Large}}) \qquad (4.9)$$

For finding the goodness-of-fit, the likelihood ratio index (ρ^2) ranges from zero, when the estimated parameters are no better than zero parameters, to one, when the estimated parameters perfectly predict the choices of the sampled decision-makers. The Pseudo-R^2 (ρ^2) of the model should be calculated which is similar to R^2 in regression. The calculation of ρ^2 is shown in equation (4.10), and it expresses the percentage of the relationship explained by the model (Koppelman and Bhat, 2006).

$$\rho^2 = 1 - \frac{\text{LL}_{\text{estimated}}}{\text{LL}_{\text{base}}} \qquad (4.10)$$

For coefficient evaluation of the variables, the T-test is used. This two-tailed test is used to test whether a particular parameter is statistically different from zero. The critical values for test statistic differ according to selected significant levels. For instance, when the significant level is 0.1, the critical values are $\pm$ 1.645. When the T-statistic of the variable is between the two critical values, the variable is insignificant, otherwise, significant. By dividing the coefficient of the variables by standard error of the variables, the T-statistic of the variables is calculated shown in equation (4.11) (Koppelman and Bhat, 2006).

$$\text{T} - \text{statistic} = \frac{\text{Coefficient}}{\text{Standard error}} \qquad (4.11)$$

4.2 Estimated Model

In this study, the multinomial logit model is used to analyze pedestrians behavior. The curbside waiting times of the pedestrian are classified into three levels. To use discrete choice model, the waiting time which is continuous data is turned into discrete data. The data is classified into three levels by determining threshold values for each of them. If another model such as regression model was used, a waiting time model could not be responded to analyze the discrete levels of waiting time to find the actual pedestrian behavior in different situations for each level of waiting time as the regression model can be used to estimate only the value of the waiting time.

In this study, three levels are considered for waiting time of pedestrians are low, medium and high levels of waiting time, and they are determined based on the level of service (LOS) of pedestrians waiting time (Nemeth et al., 2014). Table 4.1 shows the waiting time ranges at each level of waiting time. In this study, level A is considered for low level with waiting time from 0 s to 5 s, the level B is medium level for waiting time from 5 s to 10 s and the remaining levels which are C, D, E, and F are assumed as high level that includes waiting time which is equal to or higher than 10 s.

Table 4.1: Waiting time ranges at each level of waiting time, based on the LOS of pedestrians waiting time.

Level of waiting time	LOS	Comments	Waiting time ranges (s)	Level number
Low	A	Usually no conflicting traffic	0-5	1
Medium	B	Occasionally some delay due to conflicting traffic	5-10	2
High	C	Delay noticeable to pedestrians, but not inconveniencing	10-20	3
	D	Delay noticeable and irritating, increased likelihood of risk taking	20-30	
	E	Delay approaches tolerance level, risk-taking behavior likely	30-45	
	F	Delay exceeds tolerance level, high likelihood of pedestrian risk-taking	≥ 45	

In the dataset of this study, the minimum and maximum of the pedestrians waiting times are 0 s and 67.05 s, respectively.

Table 4.2 shows the data analysis of the relationship between some of the pedestrians characteristic with being at each level of waiting time.

Table 4.2: Percentage of various pedestrians' characteristic at each level of waiting time.

Passengers' characteristics		Low level (%)	Medium level (%)	High level (%)
Gender	Male	65.74	16.15	18.11
	Female	51.35	21.24	27.41
Age	Elderly	44.12	29.41	26.47
	Young	60.62	17.64	21.75
Male	Elderly	41.66	29.17	29.17
	Young	67.46	15.23	17.31
Female	Elderly	50.00	30.00	20.00
	Young	51.41	20.88	27.71

Considering the different categorizations of the sample; for the low level, the highest percentage of waiting time is recorded for young males, at the medium level is for the elderly females and at the high level is for the elderly males.

Table 4.3 displays the average of waiting time, average headway of rejected vehicles, gap, number of rejected vehicles and PWS at each level of waiting time.

Table 4.3: Average of the gap and waiting time, average headway of rejected vehicles, number of rejected vehicles and PWS at each level of waiting time.

Level of waiting time	Ranges of the waiting time (s)	The percentage share of waiting time	Average waiting time (s)	Average Gap (s)	Average headway of rejected vehicles (s)	Number of rejected vehicles	PWS
Low	0-5	59.71	1.37	6.57	0.33	0.11	1.42
Medium	5-10	18.28	7.08	4.98	2.61	1.31	1.48
High	≥ 10	22.01	19.52	6.31	3.52	5.27	1.26

The Tables 4.3 shows that the average gap at the medium level of waiting time is less than the low and high level. The PWS at the medium level is more than the low and high level.

For having a better understanding of the share of each level of waiting time, the histogram of waiting time according to each level of waiting time is presented in Figure 4.1.

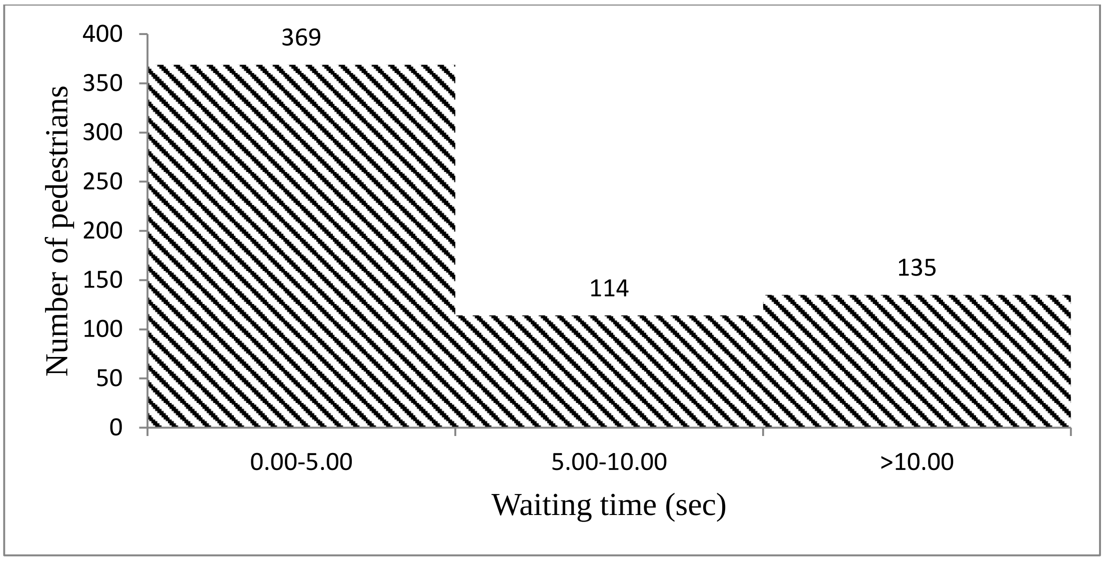

Figure 4.1: Histogram of pedestrians waiting time according to each level of waiting time.

Figure 4.1 shows that the majority of pedestrians have a low level of waiting. As it is explained, data are gathered by video recordings and traffic and pedestrians characteristics are obtained. The characteristics include six variables while five of them are independent and one is the dependent variable. In this study, the independent variables are being male and the elderly variable as dummies, the PWS variable, number of rejected vehicles variable and average headway of the rejected vehicles variable as continuous variables. The dependent variable is the pedestrians waiting time. As it is mentioned in Chapter 4 about the dummy coding, the male and the elderly variables are assumed as the dummy variable as shown in Table 4.4.

Table 4.4: Dummy coding for dummy variables.

Dummy variables	Dummy coding	
	0	1
Gender	Female	Male
Age	Young	elderly

Table 4.5 shows the correlation between variables.

Table 4.5: Existed correlation between the variables.

Variables	Gender	Age	PWS	Average headway	Number of rejected gap	Gap	Waiting time
Gender	1						-0.11
Age	0.6	1					0.11
PWS	-0.01	0.02	1				-0.09
Average headway of rejected vehicles	-0.12	0.09	-0.06	1			0.65
Number of rejected vehicles	-0.1	0.06	-0.1	0.61	1		0.93
Gap	-0.02	0.03	0.01	-0.09	0.06	1	0.02

As shown in Table 4.5, the correlation between waiting time and gap is found positive which is not logical. Hence, it is decided not to include this variable in the model. However, the correlation matrix indicates that other independent variables are appropriate to be used in the model. To have a better understanding of the relationship between the independent variables and waiting time the Figure 4.2 to Figure 4.6 are presented.

The relationship between waiting time and average headway of rejected vehicles is shown in Figure 4.2.

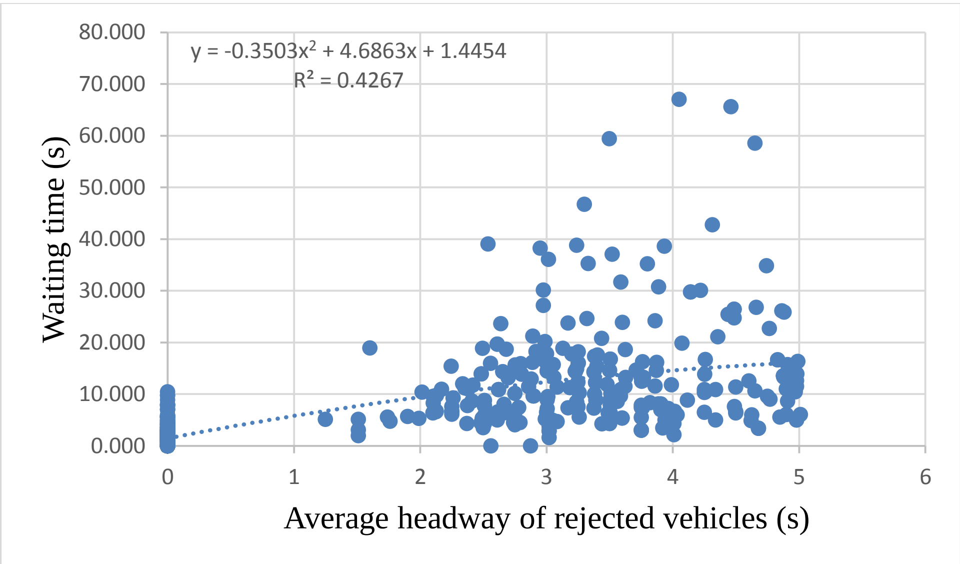

Figure 4.2: The relationship between the pedestrians waiting time and average headway of rejected vehicles.

The relationship between waiting time and number of rejected vehicles by pedestrians while waiting to cross at the curbside on the street is presented in Figure 4.3.

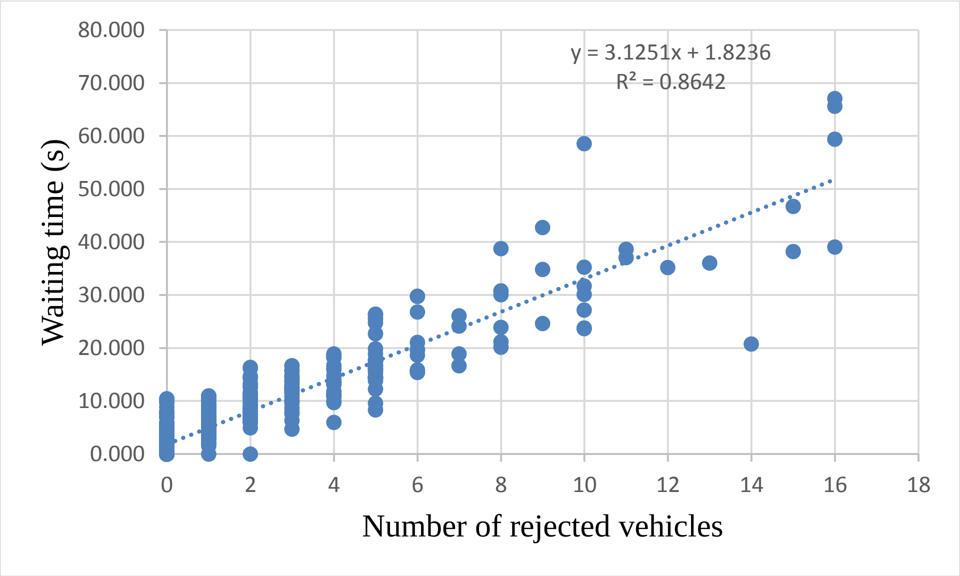

Figure 4.3: The relationship between pedestrians waiting time and the number of rejected vehicles.

The relationship between the waiting time and the number of pedestrians start waiting at the same time at the curbside on the street is displayed in Figure 4.4. The range of the number of pedestrians start waiting at the same time is between 1 and 4 pedestrians.

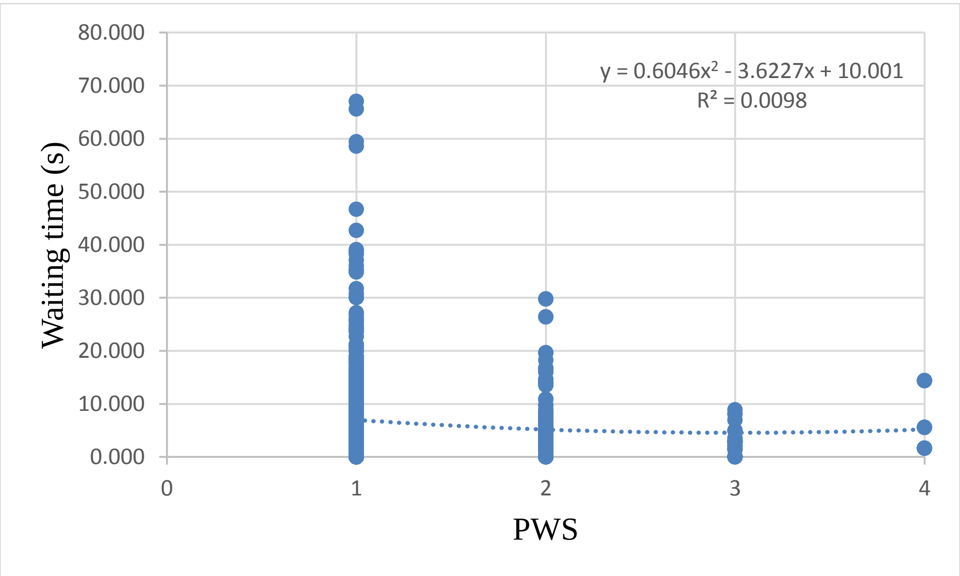

Figure 4.4: The relationship between the pedestrians waiting time and PWS.

The relation between pedestrians characteristics which are male and the elderly factors with pedestrians waiting time is presented in Figure 4.5 and Figure 4.6, respectively.

Figure 4.5: The relationship between pedestrians waiting time and male.

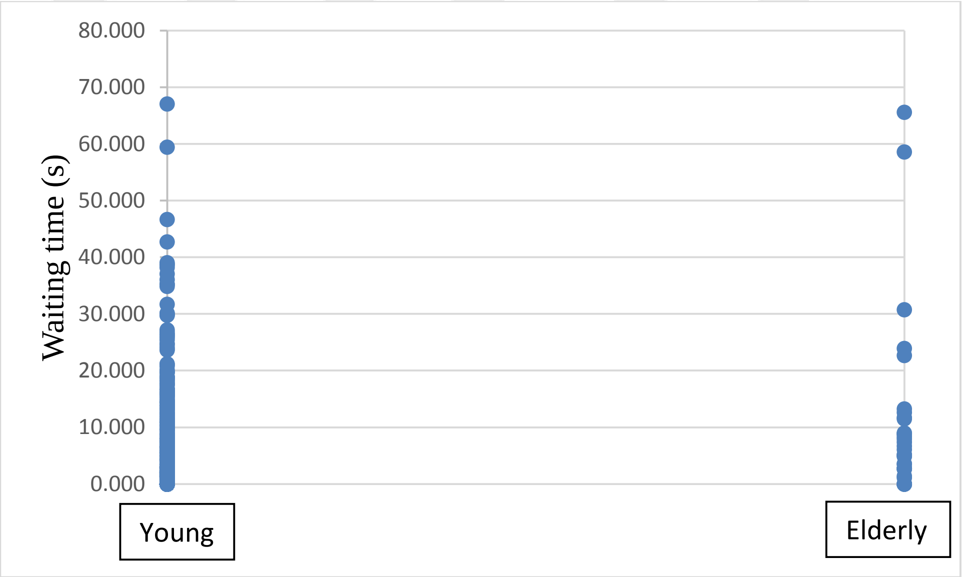

Figure 4.6: The relationship between pedestrian waiting time and age.

4.2.1 Utility equations

The high level of waiting time is selected as the reference alternative for all independent variables and constants. Hence, all the analysis and comments are relative to the high level. In this study, to have a better explanation of results obtained for the coefficient of independent variables, all variables and constants are assumed as alternative-specific. The utility equation for each level is presented in equation (4.12).

The utility of the low level: (4.12)

$$\text{Utility low} = \text{constant1} + \beta 1 \times \text{male} + \beta 2 \times \text{elderly} + \beta 3 \times \text{PWS} + \beta 4 \times \text{number of rejected vehicles} + \beta 5 \times \text{average headway of rejected vehicles}$$

The utility of medium level:

$$\text{Utility medium} = \text{constant2} + \beta 6 \times \text{male} + \beta 7 \times \text{elderly} + \beta 8 \times \text{PWS} + \beta 9 \times \text{number of rejected vehicles} + \beta 10 \times \text{average headway of rejected vehicles}$$

The utility of high level:

$$\text{Utility high} = 0$$

As it is seen in equation (4.12), the utility functions of low and medium levels have five independent variables with a constant. As a result, there are 10 alternative-specific variables and two alternative-specific constants in the model.

4.2.2 Results of the estimated model

Table 4.6 shows the coefficients and their T-statistics of the estimated model. In Table 4.6, the coefficients, which are statistically significant at 90% level, are in bold and the comparison level is the high level of waiting time, and all coefficients are in comparison to it.

Table 4.6: The estimation results for waiting time levels.

Variables	Coefficient	T-statistic
Low level		
Gender (Being male)	0.653	1.371
Age (Being elderly)	-0.109	-0.117
PWS	**0.728**	**1.920**
Average headway of rejected vehicles	**-0.996**	**-3.645**
Number of rejected vehicles	**-3.314**	**-7.239**
Constant	**7.063**	**6.252**
Medium level		
Gender (being male)	0.253	0.629
Age (being elderly)	0.409	0.533
PWS	**0.893**	**2.779**
Average headway of rejected vehicles	**-0.673**	**-2.775**
Number of rejected vehicles	**-1.447**	**-7.335**
Constant	**4.411**	**4.132**
Details		
Number of observations	618	
LL_{Base}	-588.170	
$LL_{estimated}$	-235.674	

4.2.3 Model evaluation

For evaluating the model, the probability of each level of utility for each pedestrian is calculated by equation (4.5). Then, the average of all individuals' choice probabilities for each level of waiting time is calculated. In Table 4.7, the actual shares from the observation and probabilities from the model are presented.

Table 4.7: Percentage of each level of waiting time in model and observation.

Level of waiting time	Waiting time duration at each level (s)	Model (%)	Actual share (%)
Low	0-5	59.71	59.71
Medium	5-10	18.28	18.28
High	≥10	22.01	22.01

As shown in Table 4.7, the percentage of each level of waiting time, obtained from the model is matched with the percentage of each level of waiting time gained from the observations.

As it is mentioned in Chapter 4, the likelihood ratio test is used to compare the estimated model with the base model. The LL of estimated model and base model are -235.674 and -588.170, respectively. By using equation (4.9), the -2LL is obtained as 704.992. The df is 10 which is equal to 12 coefficients of alternative-specific variables and constants in the estimated model minus 2 alternative-specific constants in the base model which is the market share model (LL(M)).

As it is shown in Figure 4.7, at the 10% level of significance, the critical chi-square value ($\chi 2$) is found as 15.99 from the chi-square table. The -2LL which is found 704.992 is higher than the critical chi-square value, so the estimated model is an improved compared to the base model.

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41

Figure 4.7: Chi-square distribution table (Scheffe, 1947).

For finding goodness-of-fit of the model, the equation (4.10) is used to calculate the ρ^2 that is equal to 0.60 which is somewhat high that represents a decent goodness-of-fit.

4.2.4 Discussion of the model findings

The results of the model show that male and the elderly variables are insignificant at both low and medium levels. These variables do not have statistical effects on waiting time in the model, but they might logically have impacts on waiting time. On the other hand, the PWS, the number of rejected vehicles and the average headway of the rejected vehicles variables are statistically significant at 90% level at both low and medium levels.

The coefficient sign of male variable is positive at both low and medium levels as expected. It means the probability of waiting time at a high level is less for males compared to females. In the literature, similarly, the waiting time for females is found to be more than the males to cross (Harrell, 1990; Tiwari et al., 2007; Ferenchak, 2016). One of the reasons is mentioned to be the higher walking speed of males (Tarawneh, 2001). Moreover, males generally make more risky crossing decisions than females and accept shorter gaps (Oxley et al., 1997; Moyano Díaz, 2002; Holland and Hill, 2010).

The coefficient sign of the elderly variable is negative at a low level. It shows that the probability of waiting time at a high level is higher than the likelihood of waiting time at a low level. However, the coefficient sign of the elderly variable is positive at the medium level which indicates that the probability of waiting time for being at the medium level for the elderly is higher than the likelihood of being at the high level. As it is mentioned in the past studies, the elderly pedestrians need more time to decide to cross the street compared to the young pedestrians (Harrell, 1990; Oxley et al., 1997; Hamed, 2001; Moyano Díaz, 2002; Li and Tsukaguchi, 2005; Holland and Hill, 2007; Rosenbloom et al., 2008; Holland and Hill, 2010; Li, 2013; Ferenchak, 2016).

The PWS variable has a positive coefficient sign at both low and medium levels of waiting time. This shows that as much as the PWS increases, the likelihood of waiting time at the high level is decreased.

The sign of the coefficient of the number of rejected variables is negative at both low and medium levels. Therefore, the higher the number of rejected variables by the pedestrians, the higher the probability of waiting time at a high level.

In this study, the average headway range of the rejected vehicles is between 1.25 s and 5.01 s. This variable has a negative sign at both low and medium levels. It is determined that in the mentioned range, as much as the average headway of the rejected vehicles is higher for the pedestrian, the probability of pedestrians waiting time at the high level is higher compared to the probability of waiting time at the low and medium levels.

5. CONCLUSIONS

The pedestrians are a vulnerable part of the transportation systems, and an in-depth understanding of their behavior is necessary to meet their needs and have a better transportation system design for their safety. The complex nature of the pedestrians behavior is affected by some factors such as individual characteristics, characteristics of the traffic flow and speed, whether the pedestrian is in a group or not.

In this thesis, a discrete choice model with five independent variables for three levels of pedestrians waiting time is estimated to have a better understanding of pedestrians crossing behavior and offer an alternative to other existing pedestrians behavior models.

The results from the discrete choice model indicate that the number of rejected vehicles variable, the PWS variable and average headway of rejected vehicles variable are the significant variables, which statistically affect the pedestrians waiting time. The male and the elderly variables are insignificant variables. However, they affect the waiting time. Moreover, the results show that if the PWS at the curbside increases, the probability of waiting at both low and medium levels is more than the probability of being at the high levels. Moreover, the higher the number of rejected vehicles by pedestrians, the higher the probability of waiting time at high levels compared to the low and medium levels. The probability of waiting time at a high level is less for males compared to females. The likelihood of waiting time at a high level is more than the probability of waiting time at a low level for elderly pedestrians. Also, the probability of waiting time at the medium level for the elderly is higher than the likelihood of waiting time at a high level. Furthermore, when the average headway of rejected vehicles which its range in this study is from 1.25s to 5.01s and 0 for pedestrians whose waiting time is 0, is higher for the pedestrian, the probability of waiting time at the high level is higher compared to the low and medium levels.

The understanding of these issues will increase the sensitivity and awareness of engineers and transport planners. This framework provides a useful guide for future pedestrians models to improve future safety.

In future studies, other methods such as the ordered probit model can be used because of the ordered nature of waiting time. The nested logit model can also be used with two composite alternatives. The first composite alternative could be allocated for pedestrians with no waiting time. Whereas, the other one is for when there is waiting time. Another possible method is for finding the estimated waiting time is the regression model. Moreover, the analysis might be expected to include other locations with different flow characteristics, number of the lane, etc. Furthermore, additional independent variables which affect waiting time that is not considered in this study can be used in future studies such as speed and type of the upcoming vehicle.

REFERENCES

Antonini, G., Bierlaire, M. and Weber, M. (2006). Discrete choice models of pedestrian walking behavior. *Transportation Research Part B*: *Methodological*, 40(8), pp.667-687.

Asaithambi, G., Kuttan, M. O. and Chandra, S. (2016). Pedestrian Road Crossing Behavior Under Mixed Traffic Conditions: A Comparative Study of an Intersection Before and After Implementing Control Measures. *Transportation in Developing Economies*. Springer International Publishing, 2(2), pp. 1–12.

Ben-Akiva, M.E. and Lerman, S.R. (1985). Discrete choice analysis: theory and application to travel demand. MIT press.

Bernhoft, I. M. and Carstensen, G. (2008). Preferences and behaviour of pedestrians and cyclists by age and gender. *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(2), pp. 83–95.

Bury, K. (1999). Statistical Distribution in Engineering. Cambridge University Press, Cambridge, UK.

Daly, A., Dekker, T. and Hess, S. (2016). Dummy coding vs effects coding for categorical variables: *Clarifications and extensions. Journal of choice modelling*, 21, pp.36-41.

Evans, D. and Norman, P. (1998). Understanding pedestrians road crossing decisions: an application of the theory of planned behaviour. *European Transport-Trasporti Europei*,13(4),pp.481–489.

Fadilah, S.I. and Shariff, A.R.M. (2014). A time gap interval for safe following distance (TGFD) in avoiding car collision in wireless vehicular networks (VANET) environment. *In Intelligent Systems, Modelling and Simulation (ISMS), 2014 5th International Conference on*, pp. 683-689. IEEE.

Ferenchak, N. N. (2016). Pedestrian age and gender in relation to crossing behavior at midblock crossings in India. *Journal of Traffic and Transportation Engineering (English Edition)*. Elsevier Ltd, 3(4), pp. 345–351.

Google Maps. (2018). Aytar Cad street. Retrivd may 01, 2018 from https://www.google.com/maps/place/Migros+MMM/@41.0749514,29.0173807,17z/data=!3m1!4b1!4m5!3m4!1s0x14cab641ce41c40d:0x2736029370d98df1!8m2!3d41.0749474!4d29.0195694.

Hamed, M. M. (2001). Analysis of pedestrians' behavior at pedestrian crossings. *Safety Science*, 38(1), pp. 63–82.

Harrell, W. (1990). Perception of risk and curb standing at street corners by older pedestrians. *Perceptual and motor skills*, 70(3 Pt 2), pp. 1363–1366.

Hensher, D.A., Rose, J.M. and Greene, W.H., (2005). Applied choice analysis: a primer. Cambridge University Press.

Holland, C. and Hill, R. (2007). The effect of age, gender and driver status on pedestrians' intentions to cross the road in risky situations. *Accident Analysis and Prevention*,39(2),pp. 224–237.

Holland, C. and Hill, R. (2010). Gender differences in factors predicting unsafe crossing decisions in adult pedestrians across the lifespan: A simulation study. *Accident Analysis and Prevention*. Elsevier Ltd, 42(4), pp. 1097–1106.

Horowitz, J.L. and Härdle, W. (1996). Direct semiparametric estimation of single-index models with discrete covariates. *Journal of the American Statistical Association*, 91(436), pp.1632-1640.

Hutchinson, P. (2008). Tailgating. Centre for Automotive Safety Research.

Iryo-Asano, M. and Alhajyaseen, W. K. M. (2014). Analysis of pedestrian clearance time at signalized crosswalks in Japan. *Procedia Computer Science*. Elsevier Masson SAS, 32(August 2014), pp. 301–308.

Ishaque, M. M. and Noland, R. B. (2008). Behavioural issues in pedestrian speed choice and street crossing behaviour: A review. *Transport Reviews*, 28(1), pp. 61–85.

Jamil, R., Xiong, S., Kong, X., Zheng, S. and Fang, Z. (2015). Pedestrian crossing patterns preference at a non-signalized crosswalk. *Procedia Manufacturing*, 3, pp.3353-3359.

Keegan, O. and O'Mahony, M. (2003). Modifying pedestrian behaviour. *Transportation Research Part A: Policy and Practice*. 37(10), pp. 889–901.

Koppelman, F. S. and Bhat, C. (2006). A Self Instructing Course in Mode Choice Modeling : Multinomial and Nested Logit Models by with technical support from Table of Contents. *Elements*, 28(3), pp. 501–12.

Lassarre, S., Papadimitriou, E., Yannis, G. and Golias, J. (2007). Measuring accident risk exposure for pedestrians in different micro-environments. *Accident Analysis & Prevention*, 39(6), pp.1226-1238.

Lee, D. N., Young, D. S. and McLaughlin, C. M. (1984). A roadside simulation of road crossing for children. *Ergonomics*, 27(12), pp. 1271–1281.

Li, B. (2013). A model of pedestrians' intended waiting times for street crossings at signalized intersections. *Transportation Research Part B: Methodological*. Elsevier Ltd, 51, pp. 17–28.

Li, Y. and Tsukaguchi, H. (2005). Relationships Between Network Topology and Pedestrian Route Choice Behavior. *The Eastern Asia Society for Transportation Studies*, 6(1), pp. 241–248.

Mitman, M. F., Cooper, D., DuBose, B., & Pande, **S**. (2010). *Driver/pedestrian Behavior at Marked and Unmarked Crosswalks in the Tahoe Basi.* (No. UCB-ITS-PRR-2010-18).

Moyano Díaz, E. (2002). Theory of planned behavior and pedestrians' intentions to violate traffic regulations. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(3), pp. 169–175.

Nemeth, B., Tillman, R., Melquist, J. and Hudson, A. (2014). Uncontrolled pedestrian crossing evaluation incorporating highway capacity manual unsignalized pedestrian crossing analysis methodology. Minnesota Department of Transportation Research Services & Library.

Okazaki, S. and Matsushita, S. (1993). A Study of Simulation Model for Pedestrian Movement with Evacuation and Queuing. *proceedings of the International Conference on Engineering for Crowd Safety*, (1), p. 428.

Oudejans, R. R., Michaels, C. F., van Dort, B., & Frissen, E. J. (1996). To cross or not to cross the effect of locomotion on street-crossing behavior. *Ecological psychology*, *8*(3), 259-267.

Oxley, J.A., Ihsen, E., Fildes, B.N., Charlton, J.L. and Day, R.H. (2005). Crossing roads safely: an experimental study of age differences in gap selection by pedestrians. *Accident Analysis & Prevention*, 37(5), pp.962-971.

Oxley, J., Fildes, B., Ihsen, E., Charlton, J. and Day, R. (1997). Differences in traffic judgements between young and old adult pedestrians. *Accident Analysis & Prevention*, 29(6), pp.839-847.

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of the association of American Geographers*, *93*(2), 314-337.

Rosenbloom, T., Ben-Eliyahu, A. and Nemrodov, D. (2008). Children's crossing behavior with an accompanying adult. *Safety Science*, 46(8), pp. 1248–1254.

Scheffe, H. (1947). The relation of control charts to analysis of variance and chi-square tests. *Journal of the American Statistical Association*, 42(239), pp.425-431.

Tarawneh, M. S. (2001). Evaluation of pedestrian speed in Jordan with investigation of some contributing factors. *Journal of Safety Research*, 32(2), pp. 229–236.

Tiwari, G., Bangdiwala, S., Saraswat, A., & Gaurav, S. (2007). Survival analysis: Pedestrian risk exposure at signalized intersections. *Transportation research part F: traffic psychology and behaviour*, *10*(2), 77-89.

World Health Organization, (2013). Pedestrian Safety: A Road Safety Manual for Decision-makers and Practitioners. World Health Organization, Geneva.

CURRICULUM VITAE

Name Surname : Ehsan AMIRNAZMIAFSHAR

Place and Date of Birth : Urmia, 05.01.1992

E-mail : amirnazmiafshar@itu.edu.tr

EDUCATION :

- **B.Sc.** : 2014, Islamic Azad University of Urmia, Civil Engineering