

Investigation of Illegal Pedestrian Crossings at Crosswalks in the Center of Athens, Greece

Marianthi Alverti ¹[0009-0002-4986-9471], Stella Roussou ¹[0009-0002-3180-8291] and George Yannis ¹[0000-0002-2196-2335]

¹ Department of Transportation Planning and Engineering – National Technical University of Athens (NTUA), Athens, Greece.

*marianthi_alverti@mail.ntua.gr, s_roussou@mail.ntua.gr, geyannis@central.ntua.gr

Abstract. Urban centers are trying to achieve safer and more sustainable mobility but understanding pedestrian behavior at crossings remains a key challenge. In cities where dense traffic and diverse road users intersect, like in the case of Athens, illegal pedestrian crossings are a persistent and hazardous phenomenon. This study explores the issue of pedestrian non-compliance, referring to the red light and jaywalking violations, at designated signalized crossings, in the center of Athens. A Stated Preference (SP) survey was structured in order to examine decision-making scenarios based on varying conditions of travel time, perceived crash risk, and personal comfort. A dataset of 212 responses was collected and analyzed by using binary logistic regression to examine the likelihood of illegal crossing, and generalized linear models to investigate the trade-offs pedestrians make between safety, time, and convenience, given three different scenarios of camera use, a traffic light with a timer, and the baseline scenario to choose from. The models highlight the influence of time pressure, traffic conditions, and social components on individual decisions. The results aim to showcase the importance of pedestrian decisions in preventing road crashes and, in this way, support more behavior-aware policy design, ultimately contributing to safer urban mobility systems.

Keywords: Pedestrian behavior, Illegal crossings, Stated Preference survey, Logistic regression, Urban mobility.

1 Introduction

Pedestrian traffic behavior is an important issue in modern cities, especially in areas with intense urban activity and high traffic density, such as the center of Athens. The center of Athens is of great interest for study as it has a great diversity in the ways and transport modes people choose to move around, such as cars, public transport, and on foot. Depending on the day and time, there are people moving to or from work, people moving for leisure and recreation, and tourists.

The illegal crossing of pedestrians, i.e., crossing outside designated crossings, crossing at red light, or in prohibited areas, is a major cause of road crashes involving

pedestrians. Pedestrian behavior is influenced by many factors, such as lack of infrastructure or poor road conditions. Examples of factors include the absence of crossings, the poor condition of pavements, and the lack of lighting. Another important factor in pedestrian behavior is low perception of the level of risk, as pedestrians may cross the road distracted or when vehicles are moving at high speed [1]. An important factor in pedestrians' decision to cross a road illegally, despite the high risk involved, is cutting down on journey time, which may often exceed their need for safety.

It is essential to correlate the data collected from the center of Athens with demographic, environmental, and behavioral characteristics of the city's pedestrians. A few studies demonstrate the complexity of pedestrian behavior and the high number of factors that influence it. Illegal crossings are not only caused by negligence or inattention but are influenced by: infrastructure factors such as waiting time at traffic signals, length of crossing, crossing design and traffic volume; socio-demographic factors, psychological factors, social factors, and driver behavior [6,9].

A questionnaire was designed to collect data through stated preference. Six scenarios were developed, which were repeated for three different types of signalers (18 in total). The scenarios had two alternatives, illegal crossing or not, for different values of travel time and crash probability. After data collection, statistical analysis was performed to produce appropriate mathematical models, which were necessary to identify the factors influencing the respondents to the alternatives given to them.

Numerous tests with various combinations of independent variables were required until the six appropriate models were developed using binomial logistic regression and one GLM model in which the factors influencing the respondents' choice of both the scenarios and the questions about their behavior were shown.

2 Methodology

2.1 Experiment

The questionnaire was created to collect data on the movement of pedestrians in the center of Athens and their characteristics and behaviors when crossing traffic lights. It was comprised of four parts and had a duration of 8-10 minutes to be completed. The questions had the main objective of being easy to comprehend in order not to confuse the respondent. The total number of questionnaires collected was 212, a sufficient number to obtain significant results from the statistical analysis. Emphasis during collection of responses was given on having a large sample in terms of age, gender, and residence, as it is important to test the behavior of pedestrians who are familiar or not with commuting in the center of Athens.

The third part of the questionnaire contained the scenarios given to the participants regarding a typical commute to the center of Athens. Participants had to choose whether or not they would cross a crossing illegally based on their time, comfort and safety. The fourth part of the questionnaire collected responses about the demographic information of the respondents.

2.2 Summary Statistics

Based on the data collected, it is clear that the respondents covered a wide age range, which is important for the interpretation of the results with the majority being young people aged 25-34 years old. It is strongly observed that the respondents consider it very important to have traffic control buttons with an indication of the remaining time on the traffic light. In the case of the remaining time indicator, the pedestrians tend to think that they know how long they will have to wait until they can cross safely and therefore have more patience.

After the collection of the questionnaire responses, all of them were coded using Microsoft Excel software. Then they were transformed into numbers to be identified by the Python Jupyter program and used for statistical analysis with GLM and Binomial logistic regression.

2.3 Statistics Theory

The method of logistic regression is used both in the development of the binary model, where the alternatives are two, and in the development of a model with more alternatives, the multinomial model. The result is a mathematical model for predicting the probability of choosing an alternative scenario, as well as describing the mode and magnitude of the influence of independent variables.

$$U_i = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n \quad (1)$$

Where U_i is the utility function of event i , $x_1 \dots x_n$, are the variables of the problem, a_0 is the constant term that shows the effect of all factors that influence the choice and have not been included as variables in the mathematical model, $a_1 \dots a_n$ are the coefficients of the variables.

The above method has been used extensively in transportation studies, in which the prediction of the influence of certain characteristics on the choice of an event is requested. The purpose of using this method is to develop a mathematical model for predicting public choice [13].

Generalized Linear Model (GLM) is a statistical regression model used when the dependent variable has more than two categories. This model examines the effect of the independent variables on the probability of an observed value in the classified dependent variable. Specifically, ordinal logistic regression is a generalization of binary logistic regression and analyzes how the independent variables affect the probability of an observed outcome belonging to one of the possible categories. This serves to understand changes in the level of the dependent variable in response to the independent variables. The GLM method is useful for a multitude of applications, such as preference evaluation, product ranking and risk assessment [14].

3 Model Development

3.1 Description of the Process

Statistical analysis involves the development of statistical models that aim to describe the choices made by respondents as a function of independent variables. The binomial regression method is based on the third part of the questionnaire, namely the questions concerning the probability that pedestrians' choice to cross illegally at a crossing would change if certain conditions, such as the presence of a control camera or a countdown indicator, were changed. The respondents' choices were either Yes or No. In contrast, the generalized linear model method is also based on the third part of the questionnaire, which contains three different scenario conditions, each containing 6 scenarios. In these scenarios, respondents were asked to choose whether they would run a red light at a crossing depending on the travel time and the probability of a crash contained in each scenario. Statistical analysis was performed using the Python programming language and the open-source application Visual Studio Code.

3.2 Models

The model building process requires that the dependent variable of each model, as well as the independent variables, be defined. The dependent variable must be binary, i.e., it must take values of 0 or 1. To find the appropriate variables that will define the models as acceptable, several tests are required, specifically 6 mathematical models were created, for each of the dependent variables: camera use, police, absence of vehicles, presence of pedestrians with the same intention, no presence of pedestrians on the same side, camera use, and countdown. Several combinations of independent variables related to gender, age, and pedestrian behavior were tested. Most of the combinations resulted in non-accepted models as they showed less than the acceptable 95% reliability.

In the case of the given scenarios in section C, the respondent was required to choose whether to cross illegally at a crossing, with different conditions for each option, such as additional waiting time, crash probability, and type of signal. Next, the variables `delta_time` and `delta_risk` were created, where the difference between the additional waiting time in each scenario and the difference in the probability of a crash are denoted. The creation of these variables is necessary to take them into account in the models and to check for their influence on the response.

Table 1. Binary Logistic Regression Results.

Variable	CAM USE		POLICING		NO CARS		ILL PED		ILL REM TI		ILL CA
	coef	pvalue	coef	pvalue	coef	pvalue	coef	pvalue	coef	pvalue	
ACCIDEN	-2.085	0.005									
ACCIDEN					-3.226	0.008					
ANY OPP					-0.712	0.019					1.317
ACCEPTA							-0.615	0.007			
ATH RES							1.386	0.005			
BUTTON	-2.860	0.003	-1.258	0.054					-1.502	0.008	
CAR MO											
CENTER			-3.222	0.001							
DIAGON			-0.915	0.001	-0.545	0.026					-1.438
DIS TRAC	1.504	0.004									
DRIVERS	-1.390	0.004							-0.560	0.013	
EDUCATI					-0.636	0.043					
EMPTY R	1.837	0.000	0.788	0.002	-0.362	0.019	-0.422	0.012			
ENCOUR	0.828	0.013									-1.971
ENTERTA	4.035	0.025									
ENTERTA	2.290	0.006									-1.675
FAM STA	-4.960	0.036					3.731	0.016	-5.715	0.000	-3.668
FASTER	0.476	0.054	0.548	0.007							
ILLEGAL	1.571	0.007	1.527	0.004							
INCOME	0.960	0.000	0.543	0.006							
IN HURR	0.942	0.007							-0.591	0.005	-1.020
JOB 2	-1.655	0.017							-1.037	0.027	
JOB 3	-5.916	0.027									
JOB 5	-3.949	0.000									3.819
LARGE U	2.060	0.000	0.508	0.034							
MORE C			-0.656	0.006							
NO OPP	-2.437	0.000	-0.747	0.004							2.003
NO VISI	-1.623	0.000									
PEDESTR									-0.524	0.019	
PHONE T	-0.756	0.037									
PUBLIC	-0.658	0.012									-0.747
RED LIG			-0.681	0.038			-0.895	0.000	-0.772	0.000	
REM TIM	3.340	0.022	2.214	0.009					-1.552	0.052	
RISK OU	-1.117	0.001	-0.750	0.005							
RISK SA					-0.478	0.027					
SHORT									0.743	0.003	
SHORTE	-1.100	0.001	-0.802	0.000							
SIGNS C							0.484	0.018			-0.811
TALK LI	0.776	0.019	0.590	0.008							
TRAF LI			-0.553	0.029			0.423	0.046			
URBAN S											-0.729
WALK C					1.324	0.036					

Table 2. Generalized Linear Model Results.

GLM Regression Scenarios			GLM Regression Scenarios			GLM Regression Scenarios		
Variable	Coef.	P> z	Variable	Coef.	P> z	Variable	Coef.	P> z
Intercept	-5,477	0,000	TRAF_LIGHT	-2,820	0,001	RISK_SAV	0,505	0,000
C(light_type) T	1,266	0,000	SHORT_DIST	0,281	0,000	SIGNS_CL	-0,123	0,030
C(light_type) T	0,883	0,000	PUBLIC_TRAN	-0,198	0,000	DIAGONAL	0,243	0,002
time	0,836	0,000	NO_CHOICE	-0,299	0,000	NO_OPPOS	0,284	0,000
risk	-0,083	0,000	SAVES_TIME	0,457	0,000	INCOME	0,119	0,007
ATH_RESIDEN	-0,608	0,001	DRIVERS_DIS	-0,536	0,000	EDUCATIO	-0,310	0,001
CENTER_ATH	-0,764	0,001	DRIVERS_SPE	0,392	0,000	CAR_MOT	0,900	0,000
WALK_CENTE	0,932	0,000	LESS_LIKELY	0,364	0,000	PROTECTI	-0,193	0,007
ACCIDENT_W	0,578	0,037	FASTER	0,284	0,000	FAM_STAT	-1,362	0,002
WEATHER	-0,880	0,000	MORE_CAUTI	-0,264	0,000	JOB_2[T.T	1,171	0,000
BUTTON	1,401	0,000	TRAF_LIGHT_	-0,320	0,000	JOB_6[T.T	1,224	0,000
REM_TIME	1,100	0,000	SHORTEST_R	0,424	0,000	WORK_ED	-0,333	0,000
WORK_EDUC	-3,395	0,000	ANY_OPPORT	-0,345	0,000	ENTERTAI	3,174	0,000
ENTERTAINM	1,865	0,000	RISK_OUTSII	-0,158	0,043	FAM_STAT	-0,792	0,000

4 Conclusion

4.1 Objectives & Methodology

An analysis of pedestrian behavior in an area with heavy traffic is necessary, as pedestrian crashes are increasing rapidly. This paper investigates illegal crossings in the center of Athens and identifies the factors that influence them. After defining the objective of the study, a literature review of related research followed to identify the characteristics of pedestrians that most influence their decision to cross illegally. After data collection, statistical analysis was performed to produce appropriate mathematical models. Through the models, it was necessary to determine the factors influencing the respondents to the alternatives given to them.

4.2 Conclusions

The conclusions that emerged from the research conducted were quite significant. It was strongly observed that a statistically significant percentage of attempts at illegal crossing occur when the traffic light is red, which confirms the importance of the problem and the need to address it immediately. Overall, the presence of a camera is the strongest deterrent to violation, as it significantly reduces offending behaviour and increases the perception of surveillance. The reasons why pedestrians tended to break the rules were mainly waiting time and the feeling that they were being delayed; the longer the waiting time, the greater the likelihood that the pedestrian would cross illegally. Another interesting observation is that residents of Athens or those who move to the city center show higher levels of infringement because they feel more familiar with their surroundings, therefore they feel more comfortable crossing illegally.

Furthermore, it has been observed by many that the sense of danger acts as a protective factor, as those who have experience of crashes and show greater compliance with the rules and caution, perhaps in fear of the consequences of their dangerous act. In fact, factors such as weather, short distance, and the absence of traffic reduce compliance with the rules, and the sense of ease that seems to exist in these circumstances may act as an excuse to break the rules. An important factor in their decision to cross illegally was also the social pressure and the intention of other pedestrians at the crossing that influence but do not always determine their final decision. As it was found, sometimes the intention of others to cross illegally pushes some people to cross imitating their behavior, while at other times it seems to act as a deterrent, projecting a more positive behavior in relation to others. Finally, it was observed that route choice reflects and biases attitudes towards the rules. Those who choose routes with traffic signals or protected crossings are less likely to break the rules, while the same does not go for those who choose the shortest route or cross diagonally where they show an increased propensity towards illegality.

4.3 Proposals for the Exploitation of the Results

Based on the results obtained from this paper, it was considered crucial to present proposals for their use in the implementation of new measures and for further research. Some suggestions for utilization of the research are to promote the importance and the need to improve the infrastructure and the conditions under which pedestrians move in the center of Athens, to reduce the rates of crashes and injuries to pedestrians, which are a burden both economically and socially for the Hellenic state. In addition, it is considered important to implement measures that will offer pedestrians safety during their movements and a case- point analysis in areas with increased pedestrian traffic, in particular the counting of pedestrians at key crossing points in order to correctly calculate green times and waiting times, a process that should be repeated frequently in areas where the volume of traffic is increasing.

Except of that, it would be crucial to organize awareness campaigns to inform citizens about the risks involved when travelling. Also, as the conclusions refer, it is necessary to redesign crossings, addition of safety islands and improve lighting and signaling with the inclusion of countdown indicators and control cameras. Finally, the worldwide use of computer vision to analyze movements, timing, and interaction with traffic lights to facilitate data analysis would be helpful too.

4.4 Suggestions for Further Research

In this paper, the behavior of pedestrians during their movement and the factors that influence their decisions on whether to cross while the traffic light is red were examined. However, there is room for further research as this is an area of concern to the scientific community worldwide. First, the sample must be expanded to include a larger part of the population that is not exclusively from the internet, and more cities in Greece or abroad have to take part in this survey, to explore possible similarities or differences and their causes.

After all the changes that will take place, the survey must be repeated at a reasonable interval to identify changes in pedestrian behavior. Also, it would be interesting if a combination of surveys for pedestrians and drivers were created to get a comprehensive picture of their interaction and how drivers are affected by the image of pedestrians. A long-term study in specific locations would help to monitor illegal crossings over time, for example, before and after infrastructure projects or social interventions. A further analysis could be done related to the estimation of the economic costs of illegal crossings through traffic delays and traffic crashes, to reinforce the need to improve road safety. Finally, it would be very helpful to conduct field mapping with the help of citizens for dangerous points with GIS tools, to personalize the analysis and improvement of the points.

References

1. Aghabayk, K., et al.: Observational-based study to explore pedestrian crossing behaviors at signalized and unsignalized crosswalks. *Accident Analysis and Prevention* 151, 105990 (2021).
2. Araya-Porras, E., Mora-Calderón, A., Aguero-Valverde, J.: Pedestrian crossing light violation in Costa Rica: exploring factors affecting mid-block crossing behavior. *Ingeniería* 32(2), 111–128 (2022).
3. Çinar, S., Yilmaz, S., Öz, B.: Pedestrians' crossing behaviors and crossing preferences: A field study. *Transactions on Transport Sciences* 13(2), 18–26 (2022).
4. Dąbrowska-Loranc, M., Sicińska, K.: Survey of pedestrians and car drivers' attitudes at the area of pedestrian crossings over period of time. *Transport Problems* 16(4), 59–70 (2021).
5. Dhoke, A., Choudhary, P.: Is there a relationship between time pressure and pedestrian non-compliance? A systematic review. *Transportation Research Part F: Traffic Psychology and Behaviour* 93, 68–89 (2023).
6. Jay, M., et al.: The light is red: Uncertainty behaviours displayed by pedestrians during illegal road crossing. *Accident Analysis and Prevention* 135, 105369 (2020).
7. Kim, E., et al.: Performance of ground-level signal detection when using a phone while walking. *Accident Analysis and Prevention* 151, 105909 (2021).
8. Ma, Y., Lu, S., Zhang, Y.: Analysis on illegal crossing behavior of pedestrians at signalized intersections based on Bayesian network. *Journal of Advanced Transportation* 2020, 2675197 (2020).
9. Mukherjee, D., Mitra, S.: A comprehensive study on factors influencing pedestrian signal violation behaviour: Experience from Kolkata City, India. *Safety Science* 124, 104610 (2020).
10. Scerri, K., Attard, M.: Understanding the impact of laziness and pedestrian intervention on car dependence within the context of a European island. *Research in Transportation Business and Management* 54, 101121 (2024).
11. Zhang, W., et al.: Analysis of pedestrian illegal crossing at unmarked segments: Environmental factors, pedestrian characteristics and crossing behaviours. *Transportation Research Part F: Traffic Psychology and Behaviour* 99, 339–355 (2023).
12. Zhu, D., Sze, N.N., Bai, L.: Roles of personal and environmental factors in the red light running propensity of pedestrian: Case study at the urban crosswalks. *Transportation Research Part F: Traffic Psychology and Behaviour* 76, 47–58 (2021).
13. Federal Highway Administration: Mathematical models for predicting mode choice. U.S. Department of Transportation (2002). Available at: <https://www.fhwa.dot.gov/policy/modalchoice/chap3.cfm>
14. GAMLj: Chapter 3 — The generalized linear model. In: *GAMLj Book* (2022). Available at: <https://gamlj.github.io/book/gzlm.html>
15. Shaaban, K., Muley, D., Mohammed, A.: Analysis of illegal pedestrian crossing behavior on a major divided arterial road. *Transportation Research Part F: Traffic Psychology and Behaviour* 54, 124–137 (2018).
16. Wickramasinghe, V., Dissanayake, S.: Evaluation of pedestrians' illegal road-crossing behavior in developing countries using conjoint analysis. (n.d.). Available at: <https://www.researchgate.net/publication/322577992>
17. Yang, Y., Sun, J.: Study on pedestrian red-time crossing behaviors: Integrated field observation and questionnaire data. (2014). Available at: <https://www.researchgate.net/publication/258256438>