Impact of Road Environment on Occurrence and Injury Severity in Hit-And-Run Crash

Bai Wei[1,2]

359954349@qq.com

[1]Department of Road Traffic Management, Sichuan Police College, Sichuan Luzhou P.O. Box 646000, China.
[2]Institute of Sichuan Police Science, Sichuan Chengdu P.O. Box 610200, China.

Abstract. In order to identify the road environment and other external causes influencing the occurrence of hit-and-run crash with casualty for enhence the road safety, and clarify the impact of these factors on injury severity in the crash, the paper employs the logistic regression model to examine external factors and crossing variables by the formation of dummy variables, based on the 2012~2014 road traffic crash data of Michigan, USA. The results indicate that nighttime, nighttime & within the lane, nighttime & weekend and rural road & nighttime have significant impacts on the occurrence of hit-and-run crash with casualty, while rural road, speed limit and rural road & within the lane are associated with the less likelihood of crash happening significantly. The well condition of the road surface, crashes occurred away from the intersection and road lanes can increase the probability of the injury severity in hit-and-run crash, while the better lighting condition, working day, daytime and urban road reduce the possibility of severe injuries happening. According to the analytic, it is necessary to strengthen the supervision of traffic at night, and more monitoring measures are needed at weekend night, rural road in evening, off-road space also.

Keywords- hit-and-run crash; road environment; logistic regression; factors identification; injury severity.

1. Introduction

The vital reason why hit-and-run behavior is regarded as one serious crime of road traffic is the troublemaker left the scene of the accident without reporting or necessary rescuing. The crashes can't be found and get effective emergency rescue in time because of irresponsible leaving and the behavior can increase the injury severity and the possibility of secondary injury in accident under special circumstances. Giving more research attention from the perspective of data mining on hit-and-run crash can not only deepen the understanding of the internal and characteristics of hit-and-run behavior, but also provide theoretical support and suggestions for improving the prevention measures of the crash.

At the present stage, studies on road traffic accident from crashes data to gain the matching characteristics and rules are carried out frequently at home and abroad[1]. Richard Tay examined the factors that might affect the occurrence of hit-and-run in fatal crashes, the gained results show that determinant factors reducing hit-and-run in fatal crashes are roadway functional class, traffic flow, types of roadway section, traffic control device, etc[2]. Kara E. MacLeod explored

the causes associated with hit-and-run pedestrian fatalities and driver identification, the conclusion of this study shows that occurrence probability of the crash increasing in the early morning, poor light conditions and on the weekend. Driver characteristics, behavior, driving history are associated with hit-and-run crash [3]. Eric Nimako Aidoo analyzed the effect of road and environmental characteristics on pedestrian hit-and-run accidents based on binary logit model. The paper pointed out the fatal accidents, unclear weather, nighttime condition can increase the likelihood of troublemaker left the scene after hitting a pedestrian significantly[4]. Chenming Jiang focused on the hit-and-run crashes occurred in urban river-crossing road tunnels and found the perpetrator's tendency to leave the crash scene without reporting is higher at night, in the tunnel exit, near to or in short tunnels[5]. Meiquan Xie explored the occurrence and injury severity of hit-and-run crash using real-time loop detector data and hierarchical Bayesian Binary Logit model with random effects [6]. Dahianna Lopez identified the influencing factors related to a hit-and-run after a vehicle-bicycle collision using logistic regression models. The paper's conclusion shows the probability of hit-and-run crash occurred partially depends on time, day of the week, and whether the vehicle type was a taxi[7]. Bei Zhou compared of factors affecting crash severities in hit-and-run and non-hit-and-run crashes based on multinomial logistic regression models. The paper indicated that compared with property damage, hit-and-run crashes were 34% more likely to result in injuries than non-hit-and-run crashes on average; and compared with property damage, hit-and-run crashes were 127% more likely to result in fatalities than non-hit-and-run crashes on average[8].

It can be known from the above analysis, there are few studies on occurrence and injury severity of hit-and-run crash from the research angle in road conditions, environment and other external factors. Besides, most of existing literatures focus on the analysis on particular condition or specific types of hit-and-run crash, but fail to make a comprehensive analysis on occurrence or injury severity of hit-and-run crash with casualty, especially in quantitative analysis. The paper employs the logistic regression model to examine external factors and crossing variables by the formation of dummy variables, based on the 2012~2014 road traffic crash data of Michigan, USA. The results indicate that nighttime, nighttime & within the lane, nighttime & weekend and rural road & nighttime have significant impacts on the occurrence of hit-and-run crash with casualty. In order to reduce the incidence of escape behaviour, it is necessary to strengthen the supervision of traffic at night, more monitoring measures are needed at weekend night, rural road in evening, off-road space also.

2. Crash Data

The Michigan Department of Transportation recorded 861651 road traffic crash data from 2012 to 2014, of which 88,771 existed hit-and-run behavior, accounting for 10.3% of the total crash data. There are 72,637 crashes accompanying with hit-and-run behavior existed in two-vehicle crash, accounting for 81.83% of the total hit-and-run crash and 14.20% of the total 511,598 two-vehicle crashes. In a total of 101,178 two-vehicle crashes causing injury or death, and 9,403 crashes involved hit-and-run behavior among them, accounting for 9.29% of the total number (Michigan Department of Transportation, 2012~2014). Take two-vehicle crashes causing injury or death as the study subject, and 90,583 crashes were selected when the troublemaker bears full responsibility for the accident and the other driver is innocent. There are 90,350 crashes are

reserved as the study sample, of which 5,791 hit-and-run crashes and accounting for 6.41% of the data sample.

The characteristics of drivers, vehicles, road environments, time and injury of the accident were recorded in detail during the study sample data, which laid a foundation for exploring the influence of road, environment and other external factors on the occurrence and the injury severity in hit-and-run crash.

It can be seen from table 1 the difference in percentage distribution on injury severity by comparing with the hit-and-run and non-hit-and-run crashes. The percentage of hit-and-run is higher than non-hit-and-run in possible injury crashes, and fewer in non-incapacitating injury crashes, incapacitating injury or fatal injury crashes. Numerical gaps between the percentage distribution of casualty are slightly (within 5%), and smaller in serious injury (within 3%). Although the minor difference exists in the overall distribution proportion of the injury severity. However, there is no doubt about the probability of damage degree in the crash will increase as the troublemaker left the scene without reporting at the first time, and the likelihood of secondary damage rise also.

Table 1. Comparison on injury severity between hit-and-run and non-hit-and-run crashes.

Damage Injury	Hit-and-run		Non-hit-and-run		Percentage Difference (%)	Dominance Ratio（hit-and-run/non-hit-and-run）
	Quantity	Percentage (%)	Quantity	Percentage (%)		
Possible Injury	4290	74.08%	58949	69.71%	4.37%	1.06
Non-incapacitating Injury	1176	20.31%	18731	22.15%	-1.84%	0.92
Incapacitating Injury or Fatal	325	5.61%	6879	8.14%	-2.52%	0.69

3. Methodology

With the development of modern statistical theory and data mining method, it is a normal practice to explore the large quantity of crash data and gain the inherent characteristics, laws and trends in road traffic accidents. As shown in table 2, the paper selects 10 factors involved external characteristics for analysis and all selected factors are calibrated in dummy variables according to the information provided by the sample crash.

Table 2. Classification of virtual variables.

Variables	Descriptions of variables	Mean	Standard Deviation
Injury severity	Possible Injury=1; Non-incapacitating Injury =2; Incapacitating Injury or Fatal =3	/	/
Intersection	Crash occurred at intersection=1; Otherwise=0	0.517	0.500
Road lane	Crash occurred in the road lane=1; Otherwise=0	0.975	0.155
Lighting	Poor lighting condition of road=1; Otherwise=0	0.113	0.317
Weather	Bad weather (snow, rain and so on)=1; Otherwise=0	0.411	0.492
Weekend	Saturday & Sunday=1; Otherwise=0	0.218	0.413
Time	Nighttime（00:00~5:00&18:00~23:00）=1; Daytime（6:00~17:00）=0	0.259	0.438
Function Classification	Rural road=1; Urban road=0	0.142	0.349
Road Surface	Poor condition of road surface (muddy, pothole and so on)=1; Otherwise=0	0.268	0.443
Speed Limit	5mph~30mph=1; 31mph~50mph=2; 51mph~120mph=3	2.036	0.662

The logistic model is one generalized linear regression model which can take the probability of different event results as the dependent variable for regression analysis, which the dependent variable of the model can be dichotomous variable or multivariate variable. The binary logistic model is used for identifying the influencing factors which affected the occurrence probability of hit-and-run crash, the model is as follows:

$$Logit(\frac{p}{1-p}) = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n \tag{1}$$

where p is the probability of crash occurred, x_n is independent variable, β_n is the coefficient of each independent variable and β_0 is one constant. When an independent variable x_n increases by one unit, with all other factors remaining constant, the odds increase by a factor $\exp(\beta_n)$ which is called the odds ratio[9]. The value of OR(odds ratio) denotes the relative amount by which the odds of the outcome increase or decrease, and the probability of hit-and-run behavior occurrence rising as OR>1 ($\beta_n > 0$) and descending when OR<1 ($\beta_n < 0$).

The ordered logistic model is used in analyzing the severity of accident widely, such as the damage degree of motorcyclist with the helmet or not [10], the severity of the crash between vehicle and bicycle and so on[11]. Considering the injury severity of the crash is classified as possible injury, non-capacitating injury, incapacitating injury or killed, and has typical characteristics of orderly selection. The ordered logistic model is employed to explore the influence on injury severity in hit-and-run crash from environment or other external factors.

The value of OR calculated from ordered logistic model indicates the odds ratio of dependent variable increase one or more grades when the independent variable increase one unit. Assuming the corresponding occurrence probability possible injury, non-capacitating injury, incapacitating injury or killed are represented as π_1, π_2 and π_3 in turn, the ordered logistic model is shown as follows:

$$Logit \frac{\pi_1}{1-\pi_1} = Logit \frac{\pi_1}{\pi_2 + \pi_3} = \alpha_1 + \beta_1 x_1 + ... + \beta_n x_n \tag{2}$$

$$Logit \frac{\pi_1 + \pi_2}{1-(\pi_1 + \pi_2)} = Logit \frac{\pi_1 + \pi_2}{\pi_3} = \alpha_2 + \beta_1 x_1 + ... + \beta_n x_n \tag{3}$$

where the α_1 and α_2 are constants, π_1, π_2 and π_3 can be concluded as follows according to the characteristics of the cumulative probability:

$$\pi_1 = \frac{\exp(-\alpha_1 + \beta_1 x_1 + ... + \beta_n x_n)}{1 + \exp(-\alpha_1 + \beta_1 x_1 + ... + \beta_n x_n)} \tag{4}$$

$$\pi_2 = \frac{\exp(-\alpha_2 + \beta_1 x_1 + ... + \beta_n x_n)}{1 + \exp(-\alpha_2 + \beta_1 x_1 + ... + \beta_n x_n)} - \pi_1 \tag{5}$$

$$\pi_3 = 1 - \pi_1 - \pi_2 \tag{6}$$

4. Results and discussion

4.1. Identification results of influencing variables in hit-and-run crash

With the aim of identifying the influencing variables which affecting the occurrence of hit-and-run crash, based on SPSS software platform, a total of 90,350 crash data sample is explored the relevance of each independent variables from Table 2 when the p-value is set 0.5. The results of univariate identification are shown in table 3:

Table3. Results of univariate identification.

Variables	Parameter Estimation	p-value	Odds Ratio
Constant	-1.669	<0.001	0.188
Nighttime	0.898	<0.001	2.454
Rural road	-0.747	<0.001	0.474
Speed limit			
31mph~50mph	-0.844	<0.001	0.430
51mph~120mph	-0.891	<0.001	0.410

There are 3 independent variables have the significant impact on the occurrence of hit-and-run crash by screening. Among the identified variables, the parameter estimation of nighttime is 0.898, and the matching odds ratio is 2.454, which stands for the occurrence probability of hit-and-run crash is higher at night. The parameter estimation and odds ratio of rural roads are -0.747 and 0.474 respectively, which shows the hit-and-run crash is less likely to occur on rural roads. The appraisal results of two factors are consistent with the actual situation. Hit-and-run behavior is less likely to be found and confirmed at night because of the poor visual field and small traffic volume, and nighttime increases the possibility of the troublemaker making a decision to leave the scene without reporting also. However, lower average traffic speed of the vehicles driving on rural roads impels the occurrence probability of hit-and-run fewer than urban roads, the small traffic volume is one important reason too. There are two comparisons in the independent variable of road speed limit, the parameter estimation of the comparison between the road speed limit at 31mph~50mph, 51mph~120mph and the speed limit range at 0mph~30mph are estimated to be -0.844, -0.891 respectively, and their odds ratio are 0.343, 0.410 in turn. Identification results indicate the trend that with the increase of road speed limit, the probability of hit-and-run behavior is smaller. Taking large traffic volume, complete road infrastructure and monitoring condition are the typical characteristic of the road with higher speed limit, which make it's easier to confirm the hit-and-run crash immediately. Therefore, with the increase of the speed limit, resulting in the less possibility of hit-and-run crash.

Considering the actual needs of hit-and-run crash prevention, with the significance level is set at 0.5 also, the identification results of cross variables which are consist of independent variable from table 2 are shown in table 4:

Table4. Identification results of cross variables.

Variables	Parameter Estimation	p-value	Odds Ratio
Constant	-2.731	<0.001	0.065
Nighttime & Crash Within the Lane	0.676	<0.001	1.967

Rural Road & Within the Lane	-1.025	<0.001	0.359	
Nighttime & weekend	0.413	<0.001	1.512	
Rural Road & Nighttime	0.492	<0.001	1.636	

There are 4 cross variables have significant influence on the occurrence of hit-and-run crash. The parameter estimation of nighttime & within the lane, nighttime & weekend, and rural road & nighttime were 0.676, 0.413, and 0.492 respectively, the corresponding odds ratio were 1.967, 1.512, and 1.636 in turn. Identification results indicate that hit-and-run crashes are more likely to occur in road lanes, weekends and rural roads when the period of time is night. The parameter estimation and odds ratio of rural road & within the lane are -1.025, 0.359 respectively, which denotes the less possibility of hit-and-run crash occurred in the rural road lanes. The confirmation results of the cross variables are related to results of univariate identification. Road lanes at night, rural roads in the evening, weekend night should be the focus in hit-and-run crash preventing and the critical factors in management measures improving.

4.2. Analysis of the influencing variables on injury severity in hit-and-run crash

In order to analyze the impact of road environment and other external factors on injury severity in hit-and-run crash, one ordered logistic regression model is used to verify 5791 crash data. The results are shown in table 5:

Table5. Calculation results for hit-and-run crashes.

Coefficient	Variables	Parameter Estimation	Standard Deviation	p-value	95% Confidence Interval	
					Lower Limit	Upper Limit
α_1	Intercept 1	0.369	0.148	0.012	0.080	0.659
α_2	Intercept 2	2.175	0.154	<0.001	1.873	2.477
β_n	Non-intersection	0.130	0.063	0.040	0.006	0.254
	Out of Road Lane	0.398	0.136	0.003	0.132	0.664
	Better Lighting	-0.300	0.085	<0.001	-0.467	-0.133
	Working Day	-0.200	0.066	0.003	-0.330	-0.070
	Daytime	-0.393	0.064	<0.001	-0.519	-0.267
	Urban road	-0.568	0.122	<0.001	-0.807	-0.330
	Well Condition of the Road Surface	0.331	0.088	<0.001	-0.159	-0.503

The parallel line test of ordinal logit regression model is 0.390, greater than 0.05, which shows the model has one good applicability.

There are 7 significant factors affecting the injury severity of hit-and-run crash through calculation. Well condition of the road surface, crashes occurred away from the intersection and road lane increase the probability of the injury severity in hit-and-run crash, and the order of influence is as follows: out of the road lane, well condition of the road surface and non-intersection. The speed of vehicles under the well condition of road surface is faster than that under the poor condition, which is the key cause in increasing the possibility of injury severity. However, the continuity of traffic in the non-intersection road area and the lower (or static) speed of vehicles outside the road lane may be one possible reason for the increasing possibility of the injury severity in hit-and-run crash. During the daytime, working days, better lighting and urban roads can lessen the possibility of increasing the damage degree in hit-and-run, and the

sequence of influence on independent variables as follows: urban roads, daytime, better lighting and working days. Visibility of the roads is higher under daytime or better lighting conditions caused the probability of hit-and-run behavior being detected in time, and emergency rescue arrives immediately, which are the important reasons to lessen the possibility of increasing injury severity in hit-and-run crash. A large number of vehicles are driven at urban roads in working days, the lower average driving speed and higher probability of being discovered in the first time after the accident are the fundamental reason which cause the decrease of the damage degree increasing possibility during the hit-and-run crash.

The vital way to lessen the degree of injury caused by hit-and-run effectively consists in complete road monitoring equipments, law-abiding driving behavior and effective emergency rescue measures. Corresponding to the identification results of significant factors affecting the occurrence of hit-and-run crash, roads with well surface condition, space out of road lanes and non-intersections area of the road should be given more concern for improving the prevention measures of hit-and-run crashes.

5. Conclusions

1, There are 3 independent variables and 4 cross variables from the external factors which influence the occurrence probability of hit-and-run crash with casualty. Nighttime, nighttime & within the lane, nighttime & weekend and rural road & nighttime have significant impacts on the occurrence of hit-and-run crash with casualty, while rural road, speed limit and rural road & within the lane are associated with the less probability of crash occurrence significantly.

2, A total of 7 significant factors affecting the injury severity of hit-and-run crash through calculation. The well condition of the road surface, crashes occurred away from the intersection and road lanes can increase the probability of the injury severity in hit-and-run crash, while the better lighting, working day, daytime and urban road lessen the possibility of severe injuries occurrence.

3, In order to reduce the occurrence of hit-and-run crash effectively, more education about traffic safety laws and rules are needed, and the road lanes at night, rural roads in the evening, weekend night, roads with well surface condition and non-intersection area should be given more focus in improving the prevention measures. In the future, it is necessary to focus on the systematic analysis about hit-and-run crash from the perspectives of environments, drivers, vehicles and so on.

Acknowledgments

References

[1] Zhou B, Li Z.Z, Zhang S.R. (2018) Comparison of factors affecting crash severities in hit-and-run and non-hit-and-run crashes. Journal of Advanced Transportation. DOI: 10.1155/2018/8537131.

[2] Jiang C.M, Lu L.J, Chen S.D, et, al. (2016) Hit-and-run Crashes in urban river-crossing road tunnels. Accident Analysis & Prevention, 95(B), 373-380.

[3] Lopez D.D, Glickman M.E, Soumerai S.B. et, al. (2018). Identifying factors related to a hit-and-run after a vehicle-bicycle collision. Journal of Transport & Health, 8, 299-306.

[4] Aidoo E.N, Amoh-Gyimah R, Ackaah W. (2013). The effect of road and environmental characteristics on pedestrian hit-and-run accidents in Ghana. Accident Analysis & Prevention, 53, 23-27.

[5] Haqverdi M Q, Seyedabrishami S, Groeger J A (2015). Identifying psychological and socio-economic factors affecting motorcycle helmet use. Accident Analysis & Prevention, 85, 102-110.

[6] Jiang X.G, Zhang G.P, Bai W et, al. (2016). Literature Review on Surrogate Exposure Technique: Quasi-induced Exposure. Journal of Southwest Jiaotong University, 51（3）, 546-552.

[7] Macleod K.E, Griswold J.B, Arnold L.S, et al. (2012). Factors associated with hit-and-run pedestrian fatalities and driver identification. Accident Analysis & Prevention, (45), 366-372.

[8] Li M, Wang Y.G, Zhang W et, al. (2015). Identifying significant factors influencing occurrence of fatal crash on mountainous freeway. China Safety Science Journal, 25(5), 126-130.

[9] Xie M.Q, Cheng W, Gill G.S et, al. (2018). Investigation of hit-and-run crash occurrence and severity using real-time loop detector data and hierarchical Bayesian Binary Logit model with random effects. Traffic Injury Prevention, 19(2): 207-213.

[10] Tay R, Barua U, Kattan L (2009). Facts contributing to hit-and-run in fatal crashes. Accident Analysis and Prevention, (41)2, 227-233.

[11] Stipancic J, Zangenehpour S, Miranda-Moreno L, et al. (2016). Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. Accident Analysis & Prevention, (97), 19-27.