WILEY | Hindawi

Research Article

Comparison of Factors Affecting Crash Severities in Hit-and-Run and Non-Hit-and-Run Crashes

Bei Zhou [iD],[1] **Zongzhi Li,**[2] **and Shengrui Zhang**[1]

[1]*School of Highway, Chang'an University, Xi'an, 710064, China*
[2]*Department of Civil, Architectural, and Environmental Engineering, Illinois Institute of Technology, Chicago, 60616, USA*

Correspondence should be addressed to Bei Zhou; bzhou3@chd.edu.cn

A hit-and-run (HR) crash occurs when the driver of the offending vehicle flees the crash scene without reporting it or aiding the victims. The current study aimed at contributing to existing literatures by comparing factors which might affect the crash severity in HR and non-hit-and-run (NHR) crashes. The data was extracted from the police-reported crash data from September 2017 to August 2018 within the City of Chicago. Two multinomial logistic regression models were established for the HR and NHR crash data, respectively. The odds ratio (OR) of each variable was used to quantify the impact of this variable on the crash severity. In both models, the property damage only (PDO) crash was selected as the reference group, and the injury and fatal crash were chosen as the comparison group. When the injury crash was taken as the comparison group, it was found that 12 variables contributed to the crash severities in both HR and NHR model. The average percentage deviation of OR for these 12 variables was 34%, indicating that compared with property damage, HR crashes were 34% more likely to result in injuries than NHR crashes on average. When fatal crashes were chosen as the comparison group, 2 variables were found to be statistically significant in both the HR and the NHR model. The average percentage deviation of OR for these 2 variables was 127%, indicating that compared with property damage, HR crashes were 127% more likely to result in fatalities than NHR crashes on average.

1. Introduction

Injuries and fatalities caused by traffic crashes are serious problems encountered by most countries in the world. According to the World Health Organization (WHO), the global number of road traffic deaths reached 1.25 million in 2013, and an additional 20-50 million were injured or disabled [1]. Among various crash types, hit-and-run (HR) crashes have drawn more and more attention in both general public and academia during past few years. A hit-and-run crash occurs when the driver of the striking vehicle leaves the crash scene without reporting it to the authority or aiding the victim [2]. According to a recent AAA (American Automobile Association) report, both HR crashes and fatalities were increasing. It was estimated that more than one HR crash happened in the US every minute in 2015 [3]. Although the penalties of HR crash vary from state to state in the US, most states consider it a felony if the crash leads to injury or fatality. HR crashes may not be completely eliminated in the short

term, but it is possible to alleviate the severe damage caused to the public by HR. In order to do so, it is crucial to understand factors affecting the severities of HR crashes. And potential engineering and administrative countermeasures could be scheduled and prioritized.

The current paper aimed at contributing to existing literatures by explicitly analyzing factors affecting the crash severities of HR crashes. Moreover, factors affecting severities of HR crashes and non-hit-and-run (NHR) crashes were quantitatively compared correspondingly.

2. Literature Review

Current literatures on HR crashes generally fall into two categories: identifying vehicles involved in HR crashes and identifying factors affecting the decisions of fleeing crash scenes. The identification of HR vehicles has been an area of interest in various fields, such as forensic, legal, and insurance [4]. For instances, Teresiński and Madro [5] examined knee

joint injuries of 357 fatal pedestrian victims of traffic crashes in order to determine the type of vehicles in HR crashes. Baucom [6] applied clinical forensic skills to identify the truck in a HR crash. Karger et al. [7] examined various fragments from the crash scene to determine if there was a preceding collision while the pedestrian was in an erect position in HR crashes.

With regard to identifying factors contributing to drivers' decisions of fleeing crash scenes, various factors have been examined in previous studies. Based on the crash data of Singapore, Tay [4] developed a binary logistic regression model to identify factors which might affect the probability of HR crashes. This study found that drivers were more likely to flee when the crash occurred at night, on straight road and near shop houses. And a crash was more likely to be a HR crash if the driver was a male, minority and aged between 45 and 69. Another study carried by Tay et al. [8] aimed at identifying factors contributing to HR decisions in fatal crashes. According to the model results, the speed limit, traffic control device, lighting condition, roadway functional class, and roadway alignment were among the most important parameters affecting the occurrence of HR in fatal crashes. Traffic engineers could target these parameters to potentially minimize the HR crashes. Regarding the most dangerous type of HR crashes, pedestrian HR crashes, Aidoo et al. [9] conducted a study to explore the effect of road and environmental characteristics on pedestrians HR crashes. Based on the results of a binary logit model, this study indicated that unclear weather, dark conditions, straight road without medians, and intersections could significantly increase the likelihood of HR decisions. MacLeod et al. [10] also explored the factors associated with HR pedestrian fatalities through logistic regression analysis. Among all factors which could increase the risk of HR, alcohol usage, and early morning were identified as the leading factors. Authors concluded that pedestrian fatality could be substantially reduced by reducing alcohol-related crashes. Aiming at comparing contributing factors in HR crashes with distracted and nondistracted drivers, Roshandeh et al. [2] conducted a comprehensive analysis based on the crash data within Cook County, Illinois. The driver distraction was classified into 5 different groups based on the distraction sources. The results of this study indicated that nondistracted drivers were 27% less likely to flee the crash scene compared to distracted drivers. Unlike most of previous studies which were carried out based on the data in developed countries, Zhang et al. [11] explored factors contributing to HR crashes using data from Guangdong Province in China. This study found that drivers who were middle-aged, male and without valid drivers' licenses were more likely to flee the scenes after crashes. Most recently, Xie et al. [12] employed real-time traffic data to investigate factors associated with HR crashes and the severity levels of HR. In addition to factors identified in previous studies, the average occupancy and speed from upstream detector were found to have a positive correlation with the occurrence of HR crashes.

As described in the literatures, most studies regarding HR crashes focused on identifying fleeing vehicles or factors contributing to the occurrence of HR crashes, but relatively few literatures discussed factors affecting the severities of HR crashes. However, studies on statistical analysis of general crash severity were relatively extensive. Savolainen et al. [13] conducted a thorough review and assessment of methodologies for highway crash injury severity analysis. A wide variety of methodological approaches were discussed, including binary outcome models, ordered discrete outcome models, unordered multinomial discrete outcome models, and other data mining approaches. This study also provided directions for future research. Considering traffic crash data were generally characterized by underreporting, Patil et al. [14] applied a weighted conditional maximum likelihood estimator to solve this problem for crash severity analysis. This new estimator was demonstrated to be able to improve the estimation quality. Targeting at injury severity analysis for a mountainous freeway section, Yu et al. [15] incorporated real-time traffic and weather data into the modelling. Three models were developed separately: fixed parameter logit model, support vector machine, and random parameter logit model. Findings from this study demonstrated that the crash injury severity could be substantially influenced by real-time traffic and weather variables. Ye et al. [16] conducted one of the first studies on the sample size requirements for crash severity modelling. Three different models were examined: multinomial logit model, ordered probit model, and mixed logit model. Results of this study confirmed that small sample size could significantly affect the quality of crash severity models. Pedestrians safety has always been a major concern in this filed. Haleem et al. [17] applied the mixed logit model to identify factors affecting pedestrian crash severity at intersections. Study results revealed that speed limit, percentage of trucks, rainy weather, and at-fault pedestrians were associated with more severe crashes at signalized intersections. A hybrid approach combining multinomial logit models and Bayesian network methods were proposed to analyze contributing factors of injury severity in rear-end crashes [18]. Based on the modelling results, several factors could significantly increase injury severities in rear-end crashes, including truck-involvement, lighting condition, windy weather, and number of vehicles involved. Naik et al. [19] investigated the relationship between weather conditions and single-vehicle truck crash severity. This study provided a practical method to combine comprehensive 15-min weather data with crash data. Similar with pedestrians, bicyclists have also been considered as vulnerable road users. Behnood et al. [20] explored factors contributing to the injury severity of bicyclists in motor-vehicle/bicycle crashes. Based on the results of a random parameters multinomial logit model, the following risk factors were identified: driver race and gender, alcohol consumption, riding on the wrong side of road, not wearing helmet, and so on. Another study conducted by Behnood et al. [21] aimed at investigating the effects of passengers on driver-injury severities in single vehicle crashes. Based on the estimation results of a random parameters logit model, the age and gender of passengers could significantly affect driver injury severities. Zeng et al. [22] proposed a generalized nonlinear model-based mixed multinomial logit approach to identify risk factors contributing to crash severity. The results indicated that the

new approach could fit the observed crash data better than the standard mixed multinomial logit model. Most recently, Jeong et al. [23] proposed a hybrid approach for classifying injury severities with imbalanced crash data. The geometric mean was used to evaluate the classification performance. The results indicated that the effect of treatments for imbalanced data was maximized when undersampling was combined with bagging training-testing method.

As mentioned before, few literatures discussed factors affecting the severities of HR crashes, let alone compared them with factors contributing to crash severities in NHR crashes. To fill this gap, the current study aimed at contributing to existing literatures by comparing factors associated with the severity level of HR and NHR crashes. The rest of this paper was organized as follows. Section 3 described the dataset and variables used in this paper. Section 4 focused on the methodology. Section 5 discussed the model results. Section 6 concluded the study and put forward the study limitations.

3. Data Collection and Processing

Data used in the current study was extracted from the police-reported crash data from September 2017 to August 2018 within City of Chicago [24]. The data contains detailed information regarding a series of crash attributes, such as "crash severity," "weather condition," and "crash type." To prepare the dataset used in the models, the original dataset was first divided into two groups: HR crashes and NHR crashes. This dataset contained 117,253 crashes, of which 30,655 were identified as HR crashes, accounting for 26.14% of total crashes. The crash severity was classified into 3 categories: fatal crash, injury crash, and property damage only (PDO) crash. As per the independent variables, 56 variables were selected for the modeling purpose. These variables fell into 11 categories, including traffic control device, device condition, weather condition, lighting condition, crash type, trafficway type, roadway surface condition, crash damage in dollar value, number of units involved (a unit refers to a motor vehicle, a pedestrian, a bicyclist, or another roadway user), crash hour, and day of week. Please refer to Table 1 for the detailed summary statistics of variables. It should be noted that some of the variables had a defined category "other/unknown," which might influence the model results. These crash entries were eliminated from the dataset before modelling. After cleaning up, the dataset contained 93,371 crashes, of which 23,332 were identified as HR crashes.

4. Methodology

In the current paper, the main objective was to identify factors which might affect the crash severities of HR crashes and NHR crashes. Over the years, various models have been proposed for the crash severity analysis based on the nature of the dependent and independent variables. In the context of this study, the dependent variable (severity of traffic crashes) was classified into three categories: property damage only (PDO) crashes, injury crashes, and fatal crashes. Due to three discrete outcomes of the dependent variable, the multinomial

logistic regression (MNL) was selected for the modelling purpose, as adopted by many previous studies in this field [25–28].

As a traditional discrete outcome model, the MNL model is suitable to analyze the relationship between potential contributing factors and multiple crash severity outcomes. As described in [29], the probability of crash n having injury severity outcome i could be written as

$$P_n(i) = P\left(O_{in} \geq O_{jn}, i, j \in I\right) \quad \forall i \neq j \tag{1}$$

where O_{in} is a function that determines the severity of crash n and I stands for a set of possible severity outcomes. Assuming O_{in} has a linear-in-parameters form, then (1) can be rewritten as

$$P_n(i) = P\left(\beta_i X_{in} + \varepsilon_{in} \geq \beta_j X_{jn} + \varepsilon_{jn}, i, j \in I\right) \quad \forall i \neq j \tag{2}$$

where β_i stands for estimated coefficients for the severity outcome i and X_{in} represents the explanatory variables which might affect the crash severity i for crash n. ε_{in} is a disturbance term which explains the unobserved influences on crash severity i. If ε_{in} is assumed to be independently and identically distributed as generalized extreme value distributed, a standard MNL model can be expressed as follows:

$$P_n(i) = \frac{\exp\left(\beta_i X_{in}\right)}{\sum_{\forall I} \exp\left(\beta_I X_{In}\right)} \tag{3}$$

Although the MNL model does not impose a monotonic effect between independent variables and dependent variables, it does require careful consideration of the correlation between dependent variables and each independent variable. Additionally, the possible multicollinearity among independent variables needs to be taken into consideration.

Before modelling, Pearson's chi-square test (χ^2) was conducted to evaluate the relationship between the dependent variable and each independent variable [30]. As a nonparametric test, Pearson's chi-square test could be used to test if two groups of categorical variables are independent of each other. The test applied a contingency table to analyze the data. To evaluate the independence, the test statistics was calculated as follows:

$$\chi^2 = \sum_{i=1}^{R} \sum_{j=1}^{C} \frac{\left(o_{ij} - e_{ij}\right)^2}{e_{ij}} \tag{4}$$

where o_{ij} and e_{ij} are the observed and expected cell count in the ith row and jth column of the table, respectively. R and C are the total number of rows and columns in the contingency table. The χ^2 value could be used to calculated the p value. Independent variables with p values greater than 0.05 were omitted from the subsequent modelling process.

To resolve the possible multicollinearity problem, the forward selection (likelihood) stepwise method was adopted in the analysis. Variables were added to the model one at a time based on the significance of the score statistic and the removal testing was based on the probability of a likelihood ratio statistic.

In the current paper, two multinomial logistic regression models were established for HR and NHR crash data, respectively. SPSS 20 was employed for the modelling purpose.

TABLE 1: Variables description and percentage distribution for HR and NHR crashes.

Variables	Descriptions of variables	HR crashes		NHR crashes	
		No. of crashes	Distribution	No. of crashes	Distribution
Crash severity					
PDO	PDO crash=1; otherwise=0	28587	93.25%	77321	89.29%
Injury	Injury crash=1; otherwise=0	2052	6.69%	9173	10.59%
Fatal	Fatal crash=1; otherwise=0	16	0.05%	104	0.12%
Traffic control device					
No control	No control=1; otherwise=0	20368	66.44%	47053	54.33%
Stop sign/flasher	Stop sign/ flasher =1; otherwise=0	2502	8.16%	9347	10.79%
Traffic signal control	Traffic signal control =1; otherwise=0	6279	20.48%	26380	30.46%
Other	other control device=1; otherwise=0	376	1.23%	1432	1.65%
Unknown	Unknown =1; otherwise=0	1130	3.69%	2386	2.76%
Device condition					
No controls	No controls=1; otherwise=0	20303	66.23%	47558	54.92%
Not functioning	Not functioning =1; otherwise=0	103	0.34%	295	0.34%
Functioning improperly	Functioning improperly =1; otherwise=0	137	0.45%	589	0.68%
Functioning properly	Functioning properly =1; otherwise=0	8083	26.37%	33204	38.34%
Other	Other condition =1; otherwise=0	205	0.67%	699	0.81%
Unknown	Unknown =1; otherwise=0	1824	5.95%	4253	4.91%
Weather					
Clear	Clear weather=1; otherwise=0	23702	77.33%	69679	80.46%
Rain	Rain =1; otherwise=0	2480	8.09%	8227	9.50%
Snow	Snow=1; otherwise=0	1236	4.03%	3538	4.09%
Cloudy	Cloudy =1; otherwise=0	759	2.48%	2550	2.94%
Other	Other weather =1; otherwise=0	189	0.62%	623	0.72%
Unknown	Unknown weather=1; otherwise=0	2286	7.46%	1981	2.29%
Lighting condition					
Daylight	Daylight=1; otherwise=0	16935	55.24%	60593	69.97%
Darkness, lighted road	Darkness, lighted road=1; otherwise=0	7970	26.00%	16929	19.55%
Darkness	Darkness=1; otherwise=0	1930	6.30%	3616	4.18%
Dusk	Dusk=1; otherwise=0	876	2.86%	2642	3.05%
Dawn	Dawn=1; otherwise=0	604	1.97%	1470	1.70%
Unknown	Unknown condition=1; otherwise=0	2340	7.63%	1348	1.56%
Crash type					
Fixed object	Fixed object=1; otherwise=0	886	2.89%	4578	5.29%
Pedestrians	Pedestrians =1; otherwise=0	972	3.17%	2196	2.54%
Parking vehicle	Parking vehicle=1; otherwise=0	13369	43.61%	12532	14.47%
Turning	Turning=1; otherwise=0	2555	8.33%	14303	16.52%
Rear-end	Rear-end=1; otherwise=0	5198	16.96%	23106	26.68%
Sideswipe same direction	Sideswipe same direction=1; otherwise=0	4128	13.47%	13895	16.05%
Sideswipe opposite direction	Sideswipe opposite direction=1; otherwise=0	472	1.54%	1278	1.48%
Pedal cyclist	Pedal cyclist=1; otherwise=0	434	1.42%	1581	1.83%

TABLE 1: Continued.

Variables	Descriptions of variables	HR crashes		NHR crashes	
		No. of crashes	Distribution	No. of crashes	Distribution
Angle	Angle=1; otherwise=0	2209	7.21%	10879	12.56%
Head on	Head on=1; otherwise=0	276	0.90%	820	0.95%
Other	Other collision type=1; otherwise=0	156	0.51%	1430	1.65%
Trafficway type					
Not divided	Not divided=1; otherwise=0	13642	44.50%	41040	47.39%
Parking lot	Parking lot=1; otherwise=0	2547	8.31%	5499	6.35%
One-way	One-way=1; otherwise=0	5910	19.28%	9300	10.74%
Divided with median barrier	Divided with median barrier=1; otherwise=0	6359	20.74%	24038	27.76%
Other	Other trafficway type=1; otherwise=0	1816	5.92%	5878	6.79%
Unknown	Unknown type=1; otherwise=0	381	1.24%	843	0.97%
Road surface condition					
Dry	Dry surface=1; otherwise=0	22386	73.03%	65549	75.69%
Wet	Wet surface=1; otherwise=0	3772	12.30%	12627	14.58%
Snow/slush	Snow/slush surface=1; otherwise=0	1357	4.43%	3435	3.97%
Other	Other condition=1; otherwise=0	219	0.71%	1010	1.17%
Unknown	Unknown condition=1; otherwise=0	2921	9.53%	3977	4.59%
Crash damage in dollar value					
$500 or less	$500 or less=1; otherwise=0	4390	14.32%	11932	13.78%
$501 - $1500	$501 - $1500=1; otherwise=0	9211	30.05%	24219	27.97%
Over $1500	Over $1500=1; otherwise=0	17054	55.63%	50447	58.25%
Number of units involved					
One unit	One unit=1; otherwise=0	908	2.96%	5571	6.43%
Two units	Two units=1; otherwise=0	27539	89.84%	75291	86.94%
More than two units	More than two units=1; otherwise=0	2208	7.20%	5736	6.62%
Day of week					
Weekday	Weekday=1; otherwise=0	21130	68.93%	64946	75.00%
Weekend	Weekend=1; otherwise=0	9525	31.07%	21652	25.00%
Crash hour					
AM peak hour	AM peak hour=1; otherwise=0	4245	13.85%	14283	16.49%
PM peak hour	PM peak hour=1; otherwise=0	5133	16.74%	15871	18.33%
Non-peak hour	Non-peak hour=1; otherwise=0	21277	69.41%	56444	65.18%

5. Model Results and Discussions

Pearson's chi-square test results for the HR and NHR model were presented in Table 2. For the HR model, the p value for "Day of Week" was greater than 0.05. And this variable was eliminated before modelling. For the NHR model, all variables were kept.

The estimation results of the multinomial logistic models for HR and NHR crashes were reported in Tables 3 and 4.

Both models fitted the data very well. The goodness of fit test results indicated deviance p values for both models were greater than 0.05. And the McFadden R^2 for HR model and NHR model were 0.358 and 0.320, respectively.

For both HR model and NHR model, PDO crashes were selected as the reference group, and the other two crash severity outcomes (injury and fatal crashes) were analyzed relative to the PDO crashes. To quantitatively analyze the impact of each variable on the crash severity outcomes, the

TABLE 2: Results of Pearson's chi-square test.

Variable	HR model		NHR model	
	Pearson'schi-square value	P value	Pearson'schi-square value	P value
Traffic control device	593.837	0.000	622.461	0.000
Device condition	546.516	0.000	564.632	0.000
Weather	101.810	0.000	152.756	0.000
Lighting condition	174.394	0.000	444.695	0.000
Crash type	9731.474	0.000	14320.314	0.000
Trafficway type	336.361	0.000	593.083	0.000
Road surface condition	81.871	0.000	182.814	0.000
Crash damage in dollar value	481.502	0.000	1460.329	0.000
Number of units involved	126.135	0.000	1176.578	0.000
Day of week	**2.293**	**0.318**	39.833	0.000
Crash hour	12.471	0.014	25.589	0.000

odds ratio (OR) of each statistically significant variable was calculated. The OR of an independent variable indicates the impact of one unit change in this variable on the comparison group relative to the reference group given other variables remain constant [31]. It can be obtained by exponentiating the coefficient of the variable. OR ranges between zero and positive infinity. For the following part, OR for variables of the HR model and NHR model were denoted as OR_{HR} and OR_{NHR}, respectively.

5.1. Injury Crashes vs. PDO Crashes. As mentioned before, PDO crashes were selected as the reference group for both HR model and NHR model. When injury crashes were taken as the comparison group, 14 and 18 variables were found to be statistically significant in the calibrated HR model and NHR model, respectively. Among these variables, 12 of them contributed to the severities of crashes in both HR and NHR model. To quantitatively compare the effects of the same variable on the crash severity in HR and NHR model, the percentage deviation of OR was calculated as follows:

$$\text{Percentage Deviation} = \frac{OR_{HR} - OR_{NHR}}{OR_{NHR}} \quad (5)$$

Please refer to Figure 1 for the detailed results.

As can be seen from Figure 1, among the 12 variables, 7 of them had larger OR in the HR model than in the NHR model. Our findings indicated that the OR of sideswipe same direction crashes in both HR and NHR models were less than 1, indicating that these crashes were more likely to be PDO crashes ($OR_{HR} = 0.327$ and $OR_{NHR} = 0.129$). However, OR_{HR} was 153% larger than OR_{NHR}, implying HR behaviors would increase the possibility of injury even in less risky crashes. The risk of injuries would be significantly increased for crashes involving pedestrians in both HR and NHR models ($OR_{HR} = 47.808$ and $OR_{NHR} = 22.037$). It should be noted that OR_{HR} was 117% larger than OR_{NHR}, indicating pedestrians were much more likely to be injured in HR crashes than in NHR crashes. Crashes involving pedal cyclist were more likely to be injury crashes than PDO crashes ($OR_{HR} = 26.562$ and $OR_{NHR} = 12.247$). Similar to pedestrians, as vulnerable road users, pedal cyclists were more likely to be injured if the offending

drivers decided to flee. Crashes occurring in parking lots were more prone to PDO crashes in both models ($OR_{HR} = 0.535$ and $OR_{NHR} = 0.305$). This could be due to relatively low speed and few people in parking lots. But again, if the drivers of the offending vehicles decided to flee, the risk of injuries would be increased by 75%. Crashes involving more than 2 units showed greater propensities towards injuries than PDO ($OR_{HR} = 6.626$ and $OR_{NHR} = 4.279$). And HR would increase the possibility of injuries by 55%. Lighting conditions might affect the crash severities in both HR and NHR models. Crashes on darkness roads with lights were more likely to be injury crashes ($OR_{HR} = 1.377$ and $OR_{NHR} = 1.333$). And the risk of injuries would be slightly increased if offending drivers fled the scenes. For crashes occurring on one-way roads, both HR and NHR model results indicated that these crashes were more likely to be PDO crashes ($OR_{HR} = 0.792$ and $OR_{NHR} = 0.789$). This could be due to the absence of opposite direction traffic on these roads. Besides, most one-ways roads are in downtown Chicago, which are strictly regulated by traffic signals. This could provide further protections to other road users, such as pedestrians.

Among the 12 variables which contributed to the severities of crashes in both HR and NHR model, 5 of them had smaller OR_{HR} than OR_{NHR}. As shown in Figure 1, the percentage deviation ranged from -2% to -71%. It should be pointed out that this did not mean HR crashes were "safer" than NHR crashes in certain conditions. For instance, the percentage deviation of OR for "parked motor vehicle" was -71%. But both OR were much smaller than 1 ($OR_{HR} = 0.056$ and $OR_{NHR} = 0.192$), indicating crashes involving parked motor vehicles were highly unlikely to result in injuries in both HR and NHR crashes.

The average percentage deviation of OR for the above 12 variables were 34%. This suggested that compared with PDO crashes, HR crashes were 34% more likely to result in injuries than NHR crashes on average.

Additionally, two variables were found to be statistically significant only in the HR model: dawn and wet road surface. HR crashes occurred at dawn were more prone to injury crashes ($OR_{HR} = 1.522$). This might because fewer witness was expected at dawn, which might delay any necessary

TABLE 3: HR multinomial logistic regression model results.

Variables	Injury crashes vs. PDO crashes			Fatal crashes vs. PDO crashes		
	Estimate	P value	Odds ratio (95% CI)	Estimate	P value	Odds ratio (95% CI)
Intercept	-4.026	0.000	-	-10.864	0.000	-
Lighting condition (relative to daylight)						
Dawn	0.420	0.048	1.522(1.003, 2.308)	-	-	-
Dusk	-	-	-	2.407	0.035	11.102(0.822, 149.864)
Darkness	-	-	-	-	-	-
Darkness, lighted road	0.320	0.000	1.377(1.193, 1.590)	1.528	0.005	4.609(0.910, 23.337)
Road surface condition (relative to dry)						
Wet	-0.265	0.004	0.767(0.641, 0.917)	-	-	-
Snow or slush	-0.547	0.005	0.579(0.395, 0.849)	-	-	-
Number of units involved (relative to one unit)						
2 units	-	-	-	-	-	-
More than 2 units	1.891	0.000	6.626(2.521, 17.414)	-	-	-
Crash type (relative to fixed object crash)						
Head on	-	-	-	-	-	-
Sideswipe opposite direction	-	-	-	-	-	-
Pedal cyclist	3.279	0.000	26.562(10.973, 64.297)	-	-	-
Pedestrian	3.867	0.000	47.808(20.042, 114.045)	-	-	-
Angle	-	-	-	-	-	-
Turning	-	-	-	-	-	-
Sideswipe same direction	-1.118	0.011	0.327(0.138, 0.776)	-	-	-
Parked motor vehicle	-2.884	0.000	0.056(0.023, 0.135)	-	-	-
Rear end	-	-	-	-	-	-
Trafficway type (relative to non-divided road)						
Divided with median barrier	0.188	0.008	1.206(1.051, 1.385)	1.603	0.025	4.970(1.224, 20.179)
One-way	-0.233	0.032	0.792(0.640, 0.980)	-	-	-
Parking lot	-0.626	0.002	0.535(0.359, 0.796)	-	-	-
Crash damage in dollar value (relative to $500 or less)						
$501 - $1500	-	-	-	-	-	-
Over $1500	1.073	0.000	2.925(2.353, 3.636)	3.368	0.002	29.024(3.466, 243.017)
Crash hour (relative to non-peak hour)						
AM peak hour	-	-	-	-	-	-
PM peak hour	-0.250	0.004	0.779(0.657, 0.923)	-	-	-

-2Log-likelihood for the model with only the intercept: 7889.860
-2Log-likelihood for the final model: 3694.060
Goodness of fit test: deviance = 2655.644, P = 1.000
Pseudo-R^2: Cox & Snell = 0.165; Nagelkerke = 0.417; McFadden = 0.358
Overall prediction accuracy: 94.8%

TABLE 4: NHR multinomial logistic regression model results.

Variables	Injury crashes vs. PDO crashes			Fatal crashes vs. PDO crashes		
	Estimate	P value	Odds ratio (95% CI)	Estimate	P value	Odds ratio (95% CI)
Intercept	-2.749	0.000	-	-7.272	0.000	-
Lighting condition (relative to daylight)						
Dawn	-	-	-	-	-	-
Dusk	-	-	-	-	-	-
Darkness	0.158	0.029	1.171(1.017, 1.350)	0.923	0.046	2.516(1.016, 6.230)
Darkness, lighted road	0.288	0.000	1.333(1.245, 1.427)	0.952	0.000	2.590(1.523, 4.407)
Road surface condition (relative to dry)						
Wet	-	-	-	-	-	-
Snow or slush	-0.342	0.000	0.710(0.610, 0.827)	-	-	-
Number of units involved (relative to one unit)						
2 units	-	-	-	-	-	-
More than 2 units	1.454	0.000	4.279 (2.952, 6.203)	1.737	0.025	5.678(1.240, 25.994)
Crash type (relative to fixed object crash)						
Head on	-	-	-	-	-	-
Sideswipe opposite direction	-0.919	0.000	0.399(0.262, 0.607)	-	-	-
Pedal cyclist	2.505	0.000	12.247(8.401, 17.853)	-	-	-
Pedestrian	3.093	0.000	22.037(15.118, 32.123)	1.986	0.015	7.288(1.474, 36.030)
Angle	-0.481	0.009	0.618(0.432, 0.885)	-2.566	0.002	0.077(0.015, 0.389)
Turning	-0.808	0.000	0.446(0.312, 0.637)	-2.870	0.001	0.057(0.011, 0.287)
Sideswipe same direction	-2.049	0.000	0.129(0.089, 0.186)	-3.863	0.000	0.021(0.003, 0.151)
Parked motor vehicle	-1.649	0.000	0.192(0.133, 0.277)	-1.927	0.010	0.146(0.033, 0.636)
Rear end	-1.286	0.000	0.276(0.193, 0.395)	-3.881	0.000	0.021(0.004, 0.115)
Trafficway type (relative to non-divided road)						
Divided with median barrier	0.211	0.000	1.235(1.165, 1.310)	-	-	-
One-way	-0.237	0.000	0.789(0.715, 0.871)	-	-	-
Parking lot	-1.189	0.000	0.305(0.253, 0.366)	-	-	-
Crash damage in dollar value (relative to $500 or less)						
$501 - $1500	-	-	-	-	-	-
Over $1500	1.225	0.000	3.405(3.035, 3.821)	2.351	0.000	10.494(3.653, 30.144)
Crash hour (relative to non-peak hour)						
AM peak hour	-	-	-	-	-	-
PM peak hour	-0.194	0.000	0.823(0.764, 0.887)	-1.099	0.019	0.333(0.133, 0.837)
Crash day of week (relative to weekday)						
Weekend	0.120	0.000	1.128(1.060, 0.199)	-	-	-

-2Log-likelihood for the model with only the intercept: 24375.043
-2Log-likelihood for the final model: 13620.463
Goodness of fit test: deviance = 9318.845, P = 1.000
Pseudo-R^2: Cox & Snell = 0.142; Nagelkerke = 0.483; McFadden = 0.320
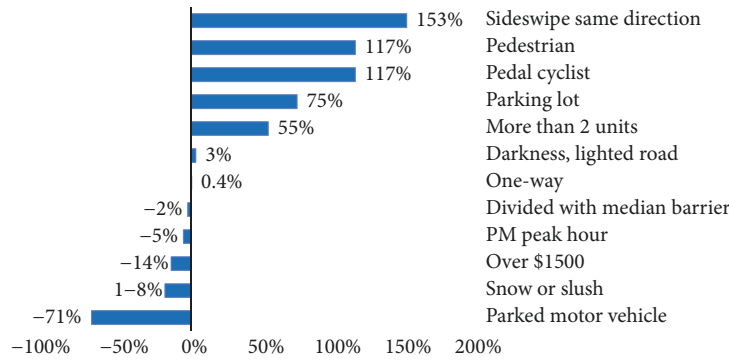Overall prediction accuracy: 90.5%

FIGURE 1: Percentage deviation of OR for statistically significant variables in HR and NHR Model (injury vs. PDO crashes).
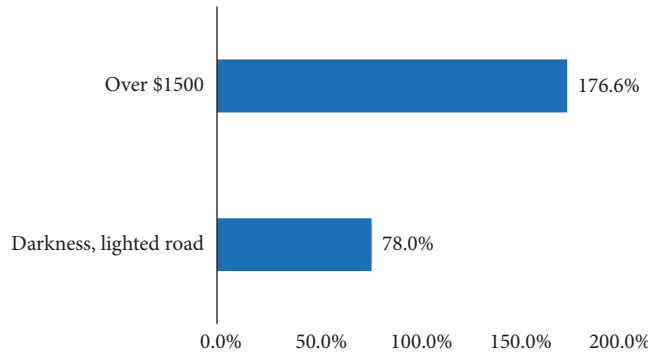

FIGURE 2: Percentage deviation of OR for statistically significant variables in HR and NHR Model (fatal vs. PDO crashes).

emergency medical service after the offending driver fled. And HR crashes on wet road were more likely to result in PDO ($OR_{HR} = 0.767$). Drivers tend to slow down on wet road surface, which might alleviate the severity of HR crashes.

Six variables were found to contribute to the crash severity only in the NHR model. Among which, 4 variables were more likely to result in PDO crashes, including "Sideswipe opposite direction" ($OR_{NHR} = 0.399$), "Angle" ($OR_{NHR} = 0.618$), "Turning" ($OR_{NHR} = 0.446$), and "Rear end" ($OR_{NHR} = 0.276$). And the presence of remaining 2 variables would more likely to lead to injury crashes, including "Darkness" ($OR_{NHR} = 1.171$) and "Weekend" ($OR_{NHR} = 1.128$)

5.2. Fatal Crashes vs. PDO Crashes. When fatal crashes were taken as the comparison group, out finding indicated that 4 and 11 variables were statistically significant in the HR and NHR model, respectively. Among which, 2 of them had potential impacts on the crash severity in both models. Similar to the previous analysis, the percentage deviations of OR for these 2 variables were presented in Figure 2.

As shown in Figure 2, both variables had significantly larger OR_{HR} than OR_{NHR}. When the crash damage value was over $1500, the possibility of fatalities would be increased considerably relative to PDO ($OR_{HR} = 29.024$ and $OR_{NHR} = 10.494$). This made sense as high damage values were normally associated with more severe crashes. It should be noted that OR_{HR} was 177% larger than OR_{NHR}, indicating that HR crashes were almost two times riskier than NHR crashes in this case. Similar to the case of injury crashes

versus PDO crashes, the lighting condition could affect the likelihood of fatalities in both HR and NHR crashes. Crashes on darkness roads with streetlights showed greater propensities towards fatalities than PDO ($OR_{HR} = 4.609$ and $OR_{NHR} = 2.59$). Again, OR_{HR} was 78% larger than OR_{NHR}, suggesting the possibility of fatality could be increased by 78% if the perpetrators fled from the scenes.

The average percentage deviation of OR for the above 2 variables was 127%. This suggested that compared with PDO crashes, HR crashes were 127% more likely to result in fatalities than NHR crashes on average.

Besides, 2 variables were found to affect the possibility of crash severity only in the HR model, including "Dusk" ($OR_{HR} = 11.102$) and "Divided with median barrier" ($OR_{HR} = 4.970$). Both OR values were greater than one, indicating that the presence of these 2 variables would considerably increase the possibility of fatalities in HR crashes. Particularly, if HR crashes occurred at dusk, the possibility of fatalities could be increased by a factor of 11.102.

On the other hand, 9 variables were found to be statistically significant only in the NHR model, including "Darkness" ($OR_{NHR} = 2.516$), "More than 2 units" ($OR_{NHR} = 5.678$), "Pedestrian" ($OR_{NHR} = 7.288$), "Angle" ($OR_{NHR} = 0.077$), "Turning" ($OR_{NHR} = 0.057$), "Sideswipe same direction" ($OR_{NHR} = 0.021$), "Parked motor vehicle" ($OR_{NHR} = 0.146$), "Rear end" ($OR_{NHR} = 0.021$), and "PM peak hour" ($OR_{NHR} = 0.333$). Most of these variables would increase the possibility of PDO crashes relative to fatal crashes with OR smaller than one. Nevertheless, the results revealed that if NHR crashes

occurred in dark environment, involving more than 2 units or pedestrians, the possibility of fatality would be increased considerably.

6. Conclusions and Recommendations

Hit-and-run (HR) crashes are those in which the driver of the offending vehicle flees the crash scene without reporting it to authorities or aiding the victims. Despite the severe punishment for HR drivers, the HR crashes and fatalities rate are still increasing in America [3]. In order to alleviate the crash severity, it is crucial to identify factors which might contribute to crash severity levels in HR crashes. Additionally, by comparing factors affecting crash severity levels in HR and NHR crashes, it would help engineers, decision-makers, and public to improve the understanding of HR crashes from a quantitative point of view.

In the current paper, the multinomial logistic regression (MLR) model was adopted to analyze the police-reported crash data from September 2017 to August 2018 within City of Chicago, Illinois. Two MLR models were established for HR crashes and NHR crashes data, respectively. In both models, PDO crashes were selected as the reference group, and injury crashes and fatal crashes were taken as the comparison group.

When injury crashes were taken as the comparison group, it was found that 12 variables contributed to the crash severity in both the HR and the NHR model. Among which, 7 variables had larger OR in the HR model than in the NHR model, including sideswipe same direction, pedestrian, pedal cyclist, parking lot, more than 2 units, darkness lighted road and one-way road. The percentage deviation of OR for these 7 variables ranged from 0.4% to 153%. The other 5 variables had smaller OR in the NHR model than in the HR model. The percentage deviation ranged from -2% to -71%. Averagely speaking, the percentage deviation for these 12 variables were 34%, indicating that compared with PDO crashes, HR crashes were 34% more likely to result in injuries than NHR crashes on average.

On the other hand, when fatal crashes were chosen as the comparison group, 2 variables were found to be statistically significant in both the HR and the NHR model, including crash damage value over $1500 and darkness lighted road. Both variables had considerably larger OR in the HR model than in the NHR model. The percentage deviations of OR for the above two variables were 177% and 78%, indicating that compared with PDO, the risk of fatality could be significantly increased if the offending driver decided to flee.

The results of the current study could help stakeholders to alleviate the HR crash severity from both the engineering and the administration perspectives. For instance, driving on dark roads with street lights could significantly increase the likelihood of fatality. Moreover, if the driver decided to flee, the likelihood of fatality could be further increased by 177%. Therefore, the traffic law enforcement again HR should be strengthened on these particular roads. Additionally, traffic safety education could be improved based on the results of the current study. For example, drivers should be aware that the risk of injury could be increased by 117% if they hit a pedestrian or pedal cyclist and decided to flee, which would substantially aggravate the punishment.

Despite the contribution of the current study, the analysis results could be further enhanced from the following aspects. The MNL model used in the current study does not impose a monotonic effect between explanatory and independent variables, but it does omit the possible unobserved effects from one severity level to the next. Future studies could benefit from applying different statistical models (such as nested logit model or mixed logit model) and comparing the results with the current study. Additionally, only one-year crash data was applied in this study. The dataset size limitation might affect the prediction accuracy, especially for the HR model. Future studies should try to use more comprehensive dataset to improve the prediction accuracy. To further capture factors affecting severities of HR and NHR crashes, the effect of additional variables should also be examined, such as alcohol consumption, driver distraction, and crash location.

Data Availability

The data used to support the findings of this study are publicly available at https://data.cityofchicago.org/Transportation/Traffic-Crashes-Crashes/85ca-t3if.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] World Health Organization, *Global Status Report on Road Safety 2015 Summary*, 2015.

[2] A. M. Roshandeh, B. Zhou, and A. Behnood, "Comparison of contributing factors in hit-and-run crashes with distracted and non-distracted drivers," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 38, pp. 22–28, 2016.

[3] W. Benson, A. Arnold, L. S. Tefft et al., *Hit-and-Run Crashes: Prevalence, Contributing Factors and Countermeasures*, AAA Foundation for Traffic Safety, 2018.

[4] R. Tay, S. M. Rifaat, and H. C. Chin, "A logistic model of the effects of roadway, environmental, vehicle, crash and driver characteristics on hit-and-run crashes," *Accident Analysis & Prevention*, vol. 40, no. 4, pp. 1330–1336, 2008.

[5] G. Teresiński and R. Madro, "Knee joint injuries as a reconstructive factors in car-to-pedestrian accidents," *Forensic Science International*, vol. 124, no. 1, pp. 74–82, 2001.

[6] B. Baucom, "Use of Clinical Forensic Skills for a 9-year-old Pedestrian Who Was Struck by a Motor Vehicle in a Hit and Run Crash," *Journal of Emergency Nursing*, vol. 32, no. 4, pp. 319–321, 2006.

[7] B. Karger, K. Teige, M. Fuchs, and B. Brinkmann, "Was the pedestrian hit in an erect position before being run over?" *Forensic Science International*, vol. 119, no. 2, pp. 217–220, 2001.

[8] R. Tay, U. Barua, and L. Kattan, "Factors contributing to hit-and-run in fatal crashes," *Accident Analysis & Prevention*, vol. 41, no. 2, pp. 227–233, 2009.

[9] E. N. Aidoo, R. Amoh-Gyimah, and W. Ackaah, "The effect of road and environmental characteristics on pedestrian hit-and-run accidents in Ghana," *Accident Analysis & Prevention*, vol. 53, pp. 23–27, 2013.

[10] K. E. MacLeod, J. B. Griswold, L. S. Arnold, and D. R. Ragland, "Factors associated with hit-and-run pedestrian fatalities and driver identification," *Accident Analysis & Prevention*, vol. 45, pp. 366–372, 2012.

[11] G. Zhang, G. Li, T. Cai, D. M. Bishai, C. Wu, and Z. Chan, "Factors contributing to hit-and-run crashes in China," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 23, pp. 113–124, 2014.

[12] M. Xie, W. Cheng, G. S. Gill, J. Zhou, X. Jia, and S. Choi, "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*, vol. 19, no. 2, pp. 207–213, 2018.

[13] P. T. Savolainen, F. L. Mannering, D. Lord, and M. A. Quddus, "The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1666–1676, 2011.

[14] S. Patil, S. R. Geedipally, and D. Lord, "Analysis of crash severities using nested logit model - Accounting for the underreporting of crashes," *Accident Analysis & Prevention*, vol. 45, pp. 646–653, 2012.

[15] R. Yu and M. Abdel-Aty, "Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data," *Safety Science*, vol. 63, pp. 50–56, 2014.

[16] F. Ye and D. Lord, "Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models," *Analytic Methods in Accident Research*, vol. 1, pp. 72–85, 2014.

[17] K. Haleem, P. Alluri, and A. Gan, "Analyzing pedestrian crash injury severity at signalized and non-signalized locations," *Accident Analysis & Prevention*, vol. 81, pp. 14–23, 2015.

[18] C. Chen, G. Zhang, R. Tarefder, J. Ma, H. Wei, and H. Guan, "A multinomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes," *Accident Analysis & Prevention*, vol. 80, pp. 76–88, 2015.

[19] B. Naik, L.-W. Tung, S. Zhao, and A. J. Khattak, "Weather impacts on single-vehicle truck crash injury severity," *Journal of Safety Research*, vol. 58, pp. 57–65, 2016.

[20] A. Behnood and F. Mannering, "Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances," *Analytic Methods in Accident Research*, vol. 16, pp. 35–47, 2017.

[21] A. Behnood and F. Mannering, "The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach," *Analytic Methods in Accident Research*, vol. 14, pp. 41–54, 2017.

[22] Z. Zeng, W. Zhu, R. Ke et al., "A generalized nonlinear model-based mixed multinomial logit approach for crash data analysis," *Accident Analysis & Prevention*, vol. 99, pp. 51–65, 2017.

[23] H. Jeong, Y. Jang, P. J. Bowman, and N. Masoud, "Classification of motor vehicle crash injury severity: A hybrid approach for imbalanced data," *Accident Analysis & Prevention*, vol. 120, pp. 250–261, 2018.

[24] Traffic Crashes - City of Chicago, https://data.cityofchicago.org/Transportation/Traffic-Crashes-Crashes/85ca-t3if.

[25] Q. Wu, G. Zhang, Y. Ci et al., "Exploratory multinomial logit model–based driver injury severity analyses for teenage and adult drivers in intersection-related crashes," *Traffic Injury Prevention*, vol. 17, no. 4, pp. 413–422, 2016.

[26] N. V. Malyshkina and F. L. Mannering, "Markov switching multinomial logit model: An application to accident-injury severities," *Accident Analysis & Prevention*, vol. 41, no. 4, pp. 829–838, 2009.

[27] S. M. Rifaat, R. Tay, and A. De Barros, "Effect of street pattern on the severity of crashes involving vulnerable road users," *Accident Analysis & Prevention*, vol. 43, no. 1, pp. 276–283, 2011.

[28] R. Tay, J. Choi, L. Kattan, and A. Khan, "A multinomial logit model of pedestrian-vehicle crash severity," *International Journal of Sustainable Transportation*, vol. 5, no. 4, pp. 233–249, 2011.

[29] S. P. Washington, M. G. Karlaftis, and F. L. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis*, CRC Press, 2011.

[30] A. Shariat-Mohaymany, A. Tavakoli-Kashani, H. Nosrati, and A. Ranjbari, "Identifying Significant Predictors of Head-on Conflicts on Two-Lane Rural Roads Using Inductive Loop Detectors Data," *Traffic Injury Prevention*, vol. 12, no. 6, pp. 636–641, 2011.

[31] A. Abdulhafedh, "Incorporating the Multinomial Logistic Regression in Vehicle Crash Severity Modeling: A Detailed Overview," *Journal of Transportation Technologies*, vol. 07, no. 03, pp. 279–303, 2017.