sustainability

MDPI

Article

Identifying the Factors Contributing to Freeway Crash Severity Based on Discrete Choice Models

Wen Cheng [1,2], Fei Ye [2,3,*], Changshuai Wang [4] and Jiping Bai [2]

[1] School of Economic and Management, Chang'an University, Xi'an 710064, China
[2] Rail Transit Department, Zhejiang Institute of Communications, Hangzhou 311112, China
[3] School of Transportation Engineering, Chang'an University, Xi'an 710064, China
[4] School of Transportation, Southeast University, Nanjing 210096, China
* Correspondence: yefei@zjvtit.edu.cn; Tel.: +86-137-3807-1989

Abstract: The freeway's operation safety has attracted wide attention. In order to mitigate the losses brought on by traffic accidents on freeways, discrete choice models were constructed based on the statistical analysis method to quantitatively analyze the significance and magnitude of the impact of multiple dimensional factors on crash severity. Based on 1154 accidents that occurred on Zhejiang Province's Hang-Jin-Qu Fressway from 2013 to 2018, the distribution characteristics of crash severity were analyzed. The dependent variable was the crash injury severity, which was categorized into property damage only (PDO), injury, and fatal. As independent variables, 15 candidate variables representing four aspects, including driver, vehicle, road, and environmental conditions, were chosen. Considering the ordered characteristics of the variables, the models developed included the ordered logit, the generalized ordered logit, and the partial proportional odds models. The Brant test found that the previous two models had difficulty dealing with the problem of partial variables that did not fit the parallel-lines assumption, and the conclusions were finally discussed through the partial proportional odds model results. The findings indicate that 11 factors have significant consequences. Five variables, namely "mountainous", "female", "driving experience 2- years", "large vehicle responsible", and "vehicle not going straight", violated the parallel-lines assumption. Female drivers and drivers aged 55+ years were more likely to suffer injuries and fatalities in collisions with guardrails and other objects. Large vehicles being involved and vehicles not going straight enhanced the likelihood of injury and fatal outcomes when drivers had 2- years of experience. Wet-skid road conditions enhanced the likelihood of injury accidents, and driving at nighttime without lighting increased the likelihood of fatal accidents. Departments responsible for traffic management can take full account of these variations and develop focused proposals for improvement.

Keywords: traffic safety; freeway crash; injury severity; discrete choice models; exploratory analysis

check for updates

Citation: Cheng, W.; Ye, F.; Wang, C.; Bai, J. Identifying the Factors Contributing to Freeway Crash Severity Based on Discrete Choice Models. *Sustainability* **2023**, *15*, 1805. https://doi.org/10.3390/su15031805

Academic Editors: Juneyoung Park, Yina Wu and Hochul Park

Received: 18 December 2022
Revised: 10 January 2023
Accepted: 11 January 2023
Published: 17 January 2023

1. Introduction

Globally, road traffic accidents kill upwards of 1.2 million people each year and injure more than 50 million people, resulting in economic losses of 3% of the global average GDP [1]. In order to alleviate the huge losses caused by road traffic accidents, efforts have been made to reduce the frequency and severity of accidents [2]. The rapid development of China's highway system has provided a strong transportation base for rapid economic growth. Unfortunately, the highways are experiencing considerable road safety problems. Among the different types of roads in China, freeways typically have the highest mortality rate [3]. The traffic fatality rate on freeways—the ratio of fatalities to injuries—is as high as 35%, significantly higher than on other types of roads [4]. The design, construction, and maintenance standards for freeway infrastructure are higher than those for other types of roads; the traffic flow is more straightforward, and crash rates may be lower [5]. However, due to the high proportion of heavy and fast-moving vehicles, freeway accidents frequently have more severe consequences [3]. According to the Traffic Administration Bureau of the Ministry of Public Security of China, freeway traffic accidents account for only 5% of road traffic accidents, but for about 10% of deaths [6]. Therefore, an in-depth study of freeway safety in China is crucial and urgently needed.

The issue of traffic safety has long been a concern for freeway management agencies and experts [7–10]. Academics have done numerous studies to examine the freeway crash frequency, which is important for developing countermeasures to reduce accidents [10–15], but not enough emphasis has been placed on the severity of the crashes. It is necessary to explore the mechanism of crash injuries and put forward countermeasures from the source by considering all relevant factors. Several intricate factors may contribute to the severity of injuries sustained in collisions [16,17]. How to quantify the impact of multiple factors on crash severity and further comprehend the interaction mechanism between elements is the premise of proposing efficient countermeasures. In recent years, there has been a gradual increase in research on crash severity in other road environments, where both discrete choice models and emerging data mining techniques have been introduced to solve crash injury severity problems [5,6,18,19]. However, only the ordered or disordered response models have been used for analysis in those using the discrete choice models [20,21]. To bridge the gap between ordered and disordered response models, the partial proportional odds model will allow some of these independent variables to violate the parallel-lines assumption [22,23]. The primary motivation of this study is to fully investigate effects of the key factors related to the driver, vehicle, road, and environmental conditions on freeway crash injury severity using a partial proportional odds model together with an ordered logit model and a generalized ordered logit model.

In order to accurately and efficiently identify the key variables that influence the severity of freeway accidents, data on 1443 historical accidents on 290 km of the Hang-Jin-Qu Freeway were collected and supplemented with accident-related road and environmental aspects. The main contribution in the research is the development of three more popular and promising discrete choice models based on historical accident data to analyze accident severity, especially the partial proportional odds model that bridges the gap between ordered and unordered models. Additionally, specific managerial recommendations were made based on the varying impacts of various elements. Finally, based on the modeling results, the potential application areas of the modeling results and the limitations of this study are discussed.

To describe the flow of the current research in detail, this article is divided into the following sections: first, the research background and main contributions of the study are presented in the introduction section; second, there is a literature review section focusing on the methods for analyzing the factors influencing the severity of accidents. Section 3 provides the databases used in this study and conducts a descriptive exploratory analysis. Section 4 presents the methods used in the study and the reasons for their selection. Then, the findings and discussion are detailed in Section 5. Lastly, Section 6 summarizes the main findings and limitations of this study.

2. Literature Review

In previous studies, a wide range of factors have been found to potentially influence the severity of road traffic crashes, including attributes of human [24,25], vehicle, road, and environment conditions [26–28].

In terms of the methodology used, the discrete choice model approach is a new trend in the literature for analyzing accident injury severity. The logit or probit models are appropriate and frequently used to solve this kind of problem [29]. Ye et al., (2013) investigated the crash frequencies by severity level for freeway sections using a joint Poisson regression model [12]. They discovered that the model could enhance the effectiveness of most coefficient estimators. Ratanavaraha and Suangka (2014) formulated a multiple logistic regression model to examine the probability of injury and fatal accidents compared with property-damage-only (PDO) accidents [27]. However, the model demands that each variable be independent and rigorously adhere to the independence from irrelevant alternative (IIA) features. Based on a random effects negative binomial (RENB) model, researchers investigated the potential factors contributing to freeway crashes [13–15]. The RENB and RPNB models significantly outperform the negative binomial (NB) model, according to research that applied a random parameters negative binomial (RPNB) model in addition [15,30]. By relaxing the IIA feature, the mixed logit models were also developed to explore the contribution of predictors of crash injury severity [31,32]. Ye et al., (2021) investigated the expressway crash severity using a random parameter logit (RPL) model by considering the potential unobserved heterogeneity [6]. Based on their investigation of these various factors' effects on safety using the RENB model, Hou, Tarko, et al., (2018) created the uncorrelated random parameter negative binomial model (URPNB) and the correlated random parameter negative binomial model (CRPNB), both of which had better goodness-of-fit [11].

To forecast the accident injury severity, several researchers have developed artificial neural network (ANN) [19], support vector machine (SVM) [18], and Markov blanket (MB) [33] models. Usually, these models provide superior model fits but are targeted at prediction accuracy and are less interpretable for accident influencing factors. By relaxing the IIA characteristics, several researchers employed the nested logit model [34] and the latent class (LC) logit model [35] to analyze the influencing factors of traffic accidents. However, because all of these models are unordered response models, they cannot capture the internal relationship between the orderly nature of some influencing factors and the injury severity.

Some academics have suggested the ordered reaction models to fit the ordered multi-classification characteristics of accident severity. Chu (2014) used ordered logit (OL) and latent class models to examine critical factors in the severity of injuries in crashes involving high-deck buses on freeways [21]. However, the OL model requires that the independent variables strictly adhere to the parallel-lines assumption (PLA); that is, the regression coefficients of the independent variables do not change with the accident's severity. Mergia et al., (2013) and Ma et al., (2016) applied a generalized ordered logit (GOL) model to quantitatively analyze the influence of the significant factors on the likelihood of crash injury severity in selected freeway areas by relaxing the PLA, which allows all independent variables to violate the assumption [28,36].

To account for the ordered nature of discrete crash severity levels and spatial association, Zeng et al., (2019) developed a Bayesian spatial GOL model with conditional autoregressive priors to assess the severity of freeway crashes. Bayesian inference shows that the spatial model outperforms the conventional GOL model because of a better model fit [20]. A partial proportional odds (PPO) model was developed by Wang et al., (2009) to evaluate the impacts of the factors and predict the injury severity in areas where freeways diverge. The results indicated that the PPO model is more adaptable and produces significantly better results [37]. The common methods used in the literature related to road traffic accident severity analysis and their advantages are shown in Table 1.

Table 1. The common methods used in the literature related to road traffic accident severity analysis.

Categories	Models	Objects	Advantages	References
Parametric predictive models	MNL	Crashes on expressways	The most commonly used	[27]
	OL	Crashes of high-deck buses on freeways	Account for the ordered nature	[21]
	GOL	Freeway crashes	Relaxing the PLA for all variables	[20,28,36]
	PPO	Freeway crashes	Allows some variables to violate the PLA	[37]
	LC	Crashes of high-deck buses on freeways	Ability to explain heterogeneity without the need to realize the form of the distribution of the assumed parameters	[21]
	RPL	Freeway crashes, E-cyclists' injury severity	Ability to explain heterogeneity	[6,38]

Table 1. *Cont.*

Categories	Models	Objects	Advantages	References
Non-parametric predictive models	ANN	Motorcyclist crashes	Facilitate the analysis of potential nonlinear relationships between variables	[19]
	SVM	Road traffic crashes	Can achieve good performance with less data	[18]
	MB	Road traffic crashes	Can select attributes by eliminating redundant variables	[33]

In conclusion, although the ordered response models can capture the ordered nature of categorical data, they impose a tight PLA on all independent variables. Some independent variables do not meet the PLA while creating the ordered response model [39]. All independent variables, however, are not constrained by the PLA in the GOL model. Both OL and GOL models lack flexibility. As a comprehensive improvement model of the OL and GOL models, the PPO model fully captures the ordered properties of each category variable and allows some variables to violate the PLA. These effects support PPO models in investigating influencing factors of freeway crash injury severity.

In terms of research contents, some studies have developed statistical models that consider several variables that can affect the severity of freeway accidents, including the driver, the vehicle, the road, and the environmental conditions [27,28]. However, not many studies have been conducted in the context of an entire freeway. The majority of the research is for a specific stretch of the freeway, such as tunnel sections [11,36,40], freeway merging and diverging locations [28]. Moreover, differences remain regarding the magnitude of the impact of various factors on accident severity. Although these problems have been paid more attention to in the modeling process in recent years, how to improve the accuracy of model prediction by enhancing the traditional discrete choice model remains to be solved.

3. Data Preparation and Description

3.1. Data Preparation

The Hang-Jin-Qu Freeway in the Zhejiang province served as this study's dataset source. The portion is a two-way, six-lane main section with a design speed of 120 km/h. We obtained the dataset from Zhejiang Traffic Group's road maintenance management system, which included details on 1443 traffic accidents over 290 km between 2013 and 2018. By excluding accidents in the service area's ramp entrance and exit or within the boundaries of the interchange, the remaining 1154 accidents with complete information were examined. The data available for each accident contained the following information: accident injuries and deaths, time and location, accident pattern and cause, vehicle and occupant characteristics, weather information, and road conditions.

Accident injuries and deaths included total number of deaths and injuries within 7 days, cause of death, and location of injury. This study did not analyze the number of specific casualties and property damage, so the casualty data were converted to accident severity levels. The severity of a traffic incident was determined by the level of injuries sustained by the most seriously injured vehicle occupant. Although there are different ways to categorize the severity of injuries, the analysis of accident severity frequently uses the traditional classification into fatal, injury, and PDO crashes [27]. Therefore, we aggregate minor and severe injuries into injury accidents. The j stands for injury severity; j = 1 denotes PDO, j = 2 denotes injury, and j = 3 denotes fatal. PDO, injury, and fatal accidents reached 424, 456, and 274, respectively, accounting for 36.7, 39.5, and 23.7% of all accidents.

Accident time records were in the form of year, month, day, hour, and week; accident location record for the freeway name, administrative level, freeway number, and specific stake number and location in the road cross-section. Vehicle and driver characteristics included the type of vehicle, nature of use and specific passenger load; driver and passenger gender, age, work, use of safety facilities, and driver age. The weather data were all recorded in real time, including sunny days and specific bad weather conditions such as rain, snow, and fog that have an important impact on accidents. Road conditions were one of the additional details included in the dataset. Road conditions included accident road alignment conditions (straight/bend/slope/curve) and road surface conditions (dry/wet/water/snow/ice). The freeway's horizontal curve properties and longitudinal profile alignment were extracted from the construction drawing design documents. The reports of the highway traffic police were used to generate data about the drivers, vehicles, roads, and environments. The data contained information about the driver's gender, age, driving history, road surface, terrain type, alignment, visibility, lighting, weather, and other factors. It is inevitable to obtain some abnormal values for reasons such as the police's record deviation when collecting accident data, or the modeling manager's artificial operation deviation. In order to increase the accuracy of the model, a distance-based outlier detection algorithm was applied to preprocess the missing values, outliers, and data consistency before statistical modeling [41]. In addition, multicollinearity between variables, i.e., correlation between two or more explanatory variables, can lead to larger standard deviations and variances of parameter estimates, and the t-test of the sample may be smaller than the critical value, thus eliminating explanatory variables that have a significant impact on the predicted value, so multicollinearity testing was performed before modeling analysis.

3.2. Data Description

The injury severity was chosen as the dependent variable. Moreover, the independent variables included drivers, road alignment, vehicle status, and environment, divided into discrete and continuous variables according to their attributes. The model can be directly modified by substituting the continuous and binary variables. Table 2 shows the independent variables' names, coding, and descriptive statistics. This study selects 15 factors for analysis.

Table 2. Description of injury-severity level frequency and percentage distribution by the explanatory variables.

Categorical Variable	Variables	Freq	Percent (%)
Injury severity	Severity = 1, PDO	424	36.7
	Severity = 2, injury	456	39.5
	Severity = 3, fatal	274	23.7
Environmental conditions	Weather condition = 0, sunny *	982	85.1
	Weather condition = 1, rainy/snowy/cloudy	172	14.9
	Visibility = 1, high visibility(200+ m) *	968	83.9
	Visibility = 2, medium visibility(100–200 m)	126	10.9
	Visibility = 3, low visibility(0–100 m)	60	5.2
	Lighting condition = 1, daylight *	675	58.5
	Lighting condition = 2, night with light	59	5.1
	Lighting condition = 3, night without light	420	36.4
Road factors	Road surface = 0, dry *	901	78.1
	Road surface = 1, wet-skid	253	21.9
	Terrain type = 0, flat *	1121	97.1
	Terrain type = 1, mountainous	33	2.9
	Vertical alignment = 0, level *	888	76.9
	Vertical alignment = 1, upgrade/downgrade	266	23.1
Driver factors	Driver gender = 0, male *	929	80.5
	Driver gender = 1, female	225	19.5
	Driver age = 1, 26–40 years *	506	43.8
	Driver age = 2, 19–25 years	94	8.2
	Driver age = 3, 41–54 years	441	38.2
	Driver age = 4, 55+ years	113	9.8
	Driving experience = 1, 3–10 years *	724	62.7
	Driving experience = 2, 2- years	121	10.5
	Driving experience = 3, 10+ years	309	26.8

Table 2. *Cont.*

Categorical Variable	Variables	Freq	Percent (%)
Vehicle factors	Collision type = 1, collision between vehicles *	869	75.3
	Collision type = 2, collision with a guardrail	169	14.6
	Collision type = 3, collision with other objects	116	10.1
	Vehicle type = 1, small vehicle *	838	72.6
	Vehicle type = 2, middle vehicle	60	5.2
	Vehicle type = 3, large vehicle	256	22.2
	Movement of vehicle = 0, going straight *	1119	97.0
	Movement of vehicle = 1, not going straight	35	3.0
Numeric variable		Mean	SD
	Radius of horizontal curve	0.170	0.462
	Length of horizontal curve	0.688	1.239
	Longitudinal gradient	0.124	0.928

Note: * The reference category; freq. = frequency; SD = standard deviation.

Figure 1 shows the frequency of accidents under various weather, visibility, and illumination conditions. Concerning the weather, the accident frequency was 5.7 times higher on sunny days than on rainy/snowy cloudy days. It was closely related to the weather along the road. According to statistics, Zhejiang Province accounts for a sizable percentage of the average yearly sunshine. As a result, 83.9% of accidents frequently occurred with good visibility. However, the distribution of accident severity proportions varied little depending on the visibility and weather. More than half of the accidents happened in daylight, and most injuries occurred during the day, according to the lighting conditions, whereas 36.4% of the accidents occurred at night without light conditions. Significant variations can also be seen in the severity levels' distribution. The frequency of accidents at night was 7.1 times higher without lighting than with lighting, which is also related to the freeway lighting distribution features. Except for some tunnels and bridges, most other sections had no additional lighting equipment.

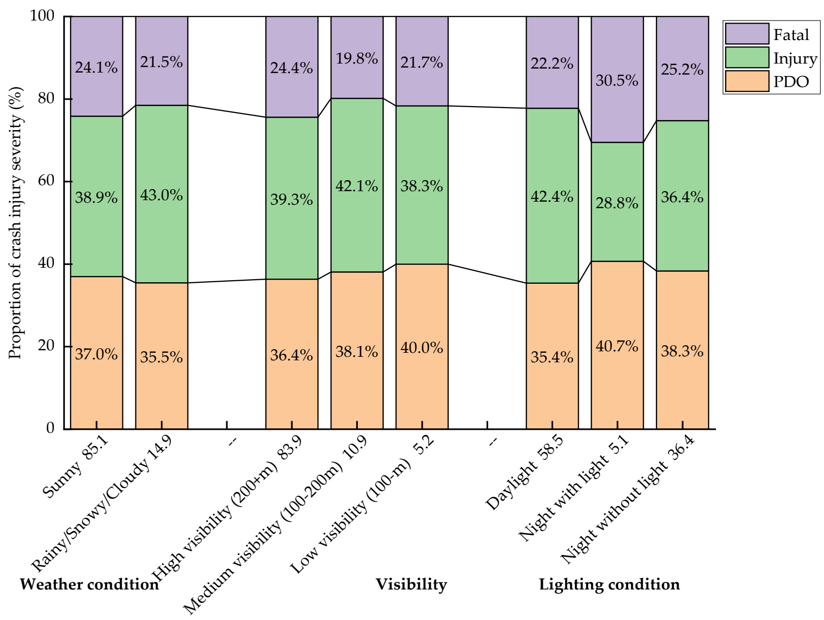

Figure 1. Injury severity in different environmental conditions.

Figure 2 shows the accident frequency distribution for road surfaces, terrain types, and vertical alignment. From the perspective of road surface, the frequency of accidents under dry road surface was 3.6 times that of wet-skid, and this distribution feature was closely related to the location's annual average weather distribution and yearly average rainfall time ratio. Under the two types of road surfaces, there were significant differences in the proportional distribution of accidents of varying severity. On the wet-skid surface, there were more injury accidents than PDO and fatal accidents, but fewer injury accidents overall. From the perspective of terrain type, most accidents occurred in flat sections, whereas just 2.9% of accidents occurred in mountainous areas. The proportion of accidents varying in severity was distributed very differently. Mountainous sections had a substantially higher percentage of injury accidents than the flat sections due to low speed limits, steady traffic flows, dense concentrations of supervision equipment, and more careful drivers. The fatal accident rate in flat sections was significantly greater than that in mountainous sections due to the higher speed limits and unstable traffic flows. According to the vertical alignment, there was a more significant difference in accident frequency between that in level and upgrade/downgrade sections, with the former being 3.3 times the latter. However, the distribution of accident proportions across different severity levels did not alter significantly.

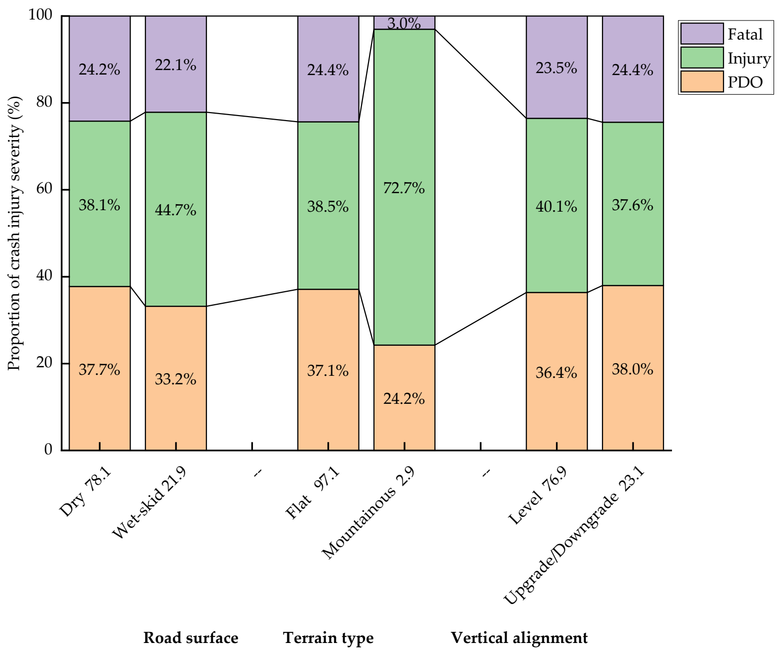

Figure 2. Injury severity in different road conditions.

Figure 3 displays the distribution of accident frequency by gender, age, and driving experience. From the perspective of gender distribution, the frequency of male accidents was 4.1 times that of female accidents, which is closely related to the distribution of the gender ratio of drivers in the area of a freeway. However, there were also significant differences in the distribution of the two in terms of severity. The proportion of injury and fatal accidents in female drivers was significantly higher than that in males. There were some variations in the severity distribution between various ages. Drivers aged 19–25 years had the highest rates of injury and fatal, possibly related to their risky driving behaviors. The severity distribution also highly varied for drivers with different levels of expertise, with drivers with 3–10 years of experience accounting for more than half of the accidents. Drivers with 3–10 years of experience had the highest percentage of injury and fatal accidents, while those with 2- years of experience had the lowest rate, which may reflect the more cautious nature of novice drivers.

The accident frequency distributions for vehicle movement, responsible vehicle type, and collision type are shown in Figure 4. There are disparities in the distribution of collision types with variable severity, with collisions between vehicles accounting for 75.3% of accidents. In contrast to collisions between vehicles, collisions with rigid guardrails and other objects were much more likely to result in injury and fatal accidents. The high speed of cars driving on the freeway was mostly to blame. The risk of injury or fatality increased dramatically in a collision between vehicles because the way the vehicles travel changed significantly. The percentage of crashes involving small vehicles was 72.6%, which is highly correlated with the distributional characteristics of vehicle type. There are apparent differences in the distribution of injury severity. The proportion of injuries and fatalities in accidents involving a small vehicle was noticeably higher. A small vehicle is more likely to cause injury accidents due to its faster speed. Accidents occurred most frequently (97% of the time) when vehicles were going straight—about 71% of collisions occurred when a vehicle was not going straight, causing PDO outcomes. The driver is generally more cautious when the vehicle is not going straight, and driving accidents also involve minor scrapes.

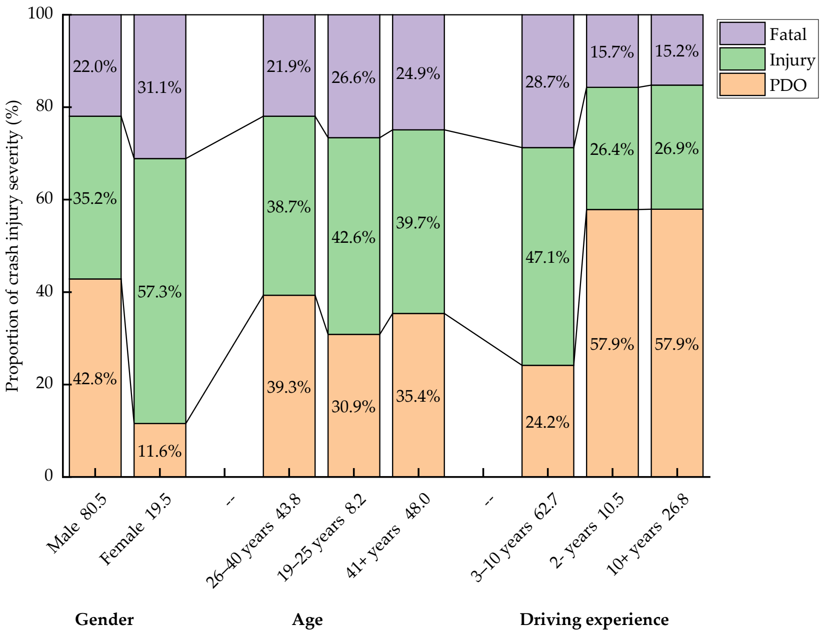

Figure 3. Injury severity according to different driver factors.

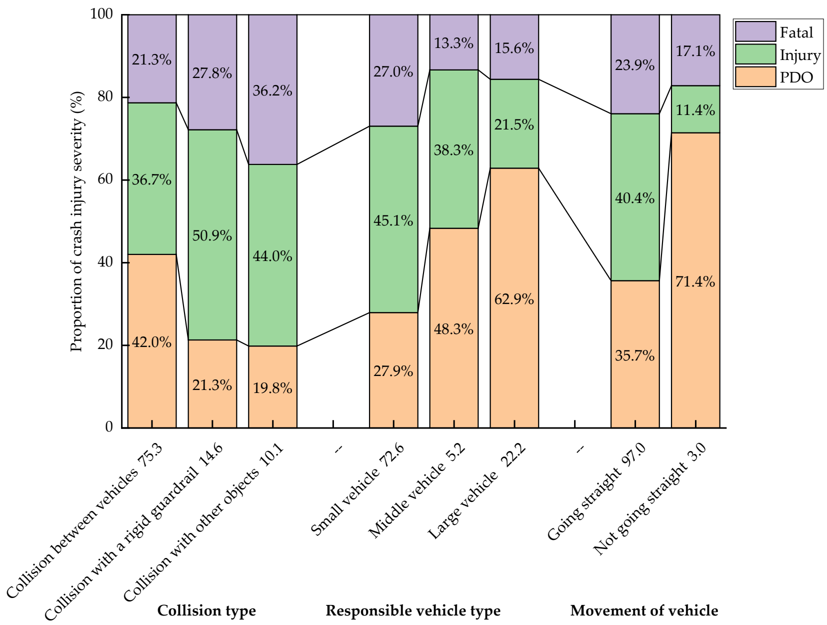

Figure 4. Injury severity according to different vehicle factors.

4. Methodology

The injury severity level is a discrete dependent variable, and a discrete choice model is an appropriate method for modeling it. The OL, GOL, and PPO models were built. Among them, the PPO model can reflect the orderly nature of each variable and allow the coefficients of some independent variables to vary with different levels, which gives it vital flexibility. The PPO is an improved model of the OL model and GOL model.

When the severity category is I ($I \geq 3$), the probability that a crash severity category i occurs in an observed crash n can be expressed as Equation (1).

$$P_n(i) = P(U_{in} \geq U_{jn}), \forall j \neq i; i, j \in I \tag{1}$$

where $P_n(i)$ is the probability that a crash severity category i occurs in an observed crash n , and U_{in} is a linear function that determines the severity of the crash n . Usually, U_{in} can be linearly expressed by Equation (2).

$$U_{in} = \beta_i X_n + \varepsilon_{in} \tag{2}$$

where X_n is a vector of measurable characteristics (risk factors) that determine severity, β_i is a vector of computable coefficients to be estimated, and ε_{in} is a disturbance term that considers unobserved effects.

Here, we define $j = 1$ as the lowest value of the injury severity variable, i.e., PDO. An ordered discrete choice model is an appropriate method for modeling it [20]. The OL model probability can be expressed as Equation (3).

$$P(Y_i > j) = \frac{\exp(\alpha_j + \mathbf{X\beta})}{1 + \exp(\alpha_j + \mathbf{X\beta})}, \quad j = 1, 2, \cdots, M - 1 \tag{3}$$

where $P(Y_i > j)$ represents the probability of crash severity for a given accident i. j is the number of cut points. α_j represents the regression intercept of each cut point. $\boldsymbol{\beta}$ is the regression coefficient vector that does not change across different logit models. $\mathbf{X}$ is the explanatory variables vector.

However, a strict limitation of using the OL model is the PLA [42]. Therefore, some scholars put forward the GOL model as an alternative method that can relax the PLA's limitations. The only difference between them is that the regression coefficients β_j may differ in severity levels.

The probability calculation expression of GOL is as Equation (4).

$$P(Y_i > j) = \frac{\exp(\alpha_j + X_i\beta_j)}{1 + \exp(\alpha_j + X_i\beta_j)}, \quad j = 1, 2, \cdots, M - 1 \tag{4}$$

where j represents an injury severity category, α_j is the cutoff point for the jth cumulative logit, β_j is a vector of model parameters, and X_i is a vector of observed explanatory variables [43].

In practice, one or several variables may violate the PLA [42]. In such a situation, some of the parameters of variables satisfying the PLA may be redundant [43]. Hence, a gamma-parameterized form of the GOL model proposed by Peterson and Harrell (1990) is commonly used [23]. The gamma-parameterized GOL model is commonly known as the partially constrained GOL or PPO model. The PPO model is the intermediate method between the OL and GOL models [22,23]. In the PPO model, the PLA is only relaxed for some variables. In other words, the regression coefficients of explanatory variables that violate the PLA vary across the dividing points, while other variables remain unchanged. The PPO model can be described as Equation (5).

$$P(Y > j) = \frac{\exp[\alpha_j + (X_i\beta_j + T_i\gamma_j)]}{1 + \exp[\alpha_j + (X_i\beta_j + T_i\gamma_j)]}, \quad j = 1, 2 \tag{5}$$

where T_i represents a subset of explanatory variables for which the PLA is violated and γ_j is a vector of parameters associated with T_i [42]. The γ_j represent deviations from proportionality. If all gammas are equal to zero, the model reduces to the traditional OL model. The model's parameters are estimated using the maximum likelihood procedure [43]. The probability formula of the PPO model can be expressed as Equations (6)–(8).

$$P(Y = 1 \mid \mathbf{X}) = 1 - g(\mathbf{X}\beta_1) \tag{6}$$

$$P(Y = 2 \mid \mathbf{X}) = g(\mathbf{X}\beta_1) - g(\mathbf{X}\beta_2) \tag{7}$$

$$P(Y = 3 \mid \mathbf{X}) = g(\mathbf{X}\beta_2) \tag{8}$$

The PPO model can fit the command "gologit2 with autofit lrf" written by the user in the Stata 16.0 [44]. The PPO model results are interpreted similarly to a binary logistic regression [43]. We can group the three outcome levels into two comparison groups for a variable with three categories. Consequently, this results in two sets of outcome groups for each model developed. For $j = 1$, outcome level 1 is compared with outcome levels 2 and 3; for $j = 2$, the comparison is between outcome levels 1 and 2 compared with outcome level 3. An optimistic coefficient indicates that higher values on the predictor variable increase the likelihood of an injury being at a more severe level than the current one. Likewise, a negative coefficient indicates that higher values on the predictor variable increase the likelihood of an injury being in the present or lower level; that is, it decreases the likelihood of being in higher injury groups.

The marginal effects provide the effect of a one-unit increase in an explanatory variable on the injury-outcome probability. The average marginal effects of overall crash observations will be computed and reported to assess the influence of the explanatory variables on injury severity outcome probabilities. For the crash i and injury severity level j, the marginal effects of all variables can be expressed as Equation (9).

$$E^{P_{ij}}_{X_{ijk}} = P_{ij}(X_{ijk} = 1) - P_{ij}(X_{ijk} = 0) \tag{9}$$

where $E^{P_{ij}}_{X_{ijk}}$ represents the marginal effects of the kth dummy variable X_{ijk}, and $P_{ij}(X_{ijk} = 1)$ and $P_{ij}(X_{ijk} = 0)$ denote the probability that the dummy variable X_{ijk} equals 1 and 0, respectively.

It is usually necessary to test the model's validity in a regression analysis. The Akaike information criterion (AIC), Bayesian information criterion (BIC), and McFadden's pseudo R^2 are usually used to evaluate the fitness of a theoretical model. The calculation method of each evaluation index can be defined as Equations (10)–(12).

$$AIC = -2LL(\beta) + 2K \tag{10}$$

$$BIC = -2LL(\beta) + \ln(N) \times K \tag{11}$$

$$PseudoR^2 = 1 - \frac{LL(\beta)}{LL(0)} \tag{12}$$

where $LL(0)$ is the initial value of the log-likelihood at zero; that is, the value of the log-likelihood when no independent variable is included in the model. $LL(\beta)$ is the convergence value of the log-likelihood function; that is, the value of the log-likelihood function when all significant independent variables and constant terms are included in the model, K is the number of parameters, and N is the number of observations. The best-fitted models have a lower AIC, BIC, and higher pseudo R^2.

5. Results and Discussion

5.1. Model Parameter Estimation

Table 3 shows the parameter estimation results of the OL model, which only include the statistically significant variables (90% significance level). The collision type, road surface, lighting condition, terrain type, driver gender, driver age, driving experience, vehicle type, and movement of vehicle all affect the injury severity. Consequently, a positive sign for a variable means that, compared to crashes that correspond to the feature represented by the variable's base category, crashes with the variable's characteristic are more likely to be assigned to the high injury level relative to the base segment. The "collision type" results indicate an increased likelihood of injury severity being transferred to a high level in the case that the vehicle collides with a guardrail or other objects compared to a collision between vehicles. These results are in line with expectations and are generally consistent with the findings of previous studies on the analysis of factors influencing the severity of freeway crashes.

Table 3. Coefficients (and standard errors) of the OL model for injury-severity outcomes.

Variables	Description	Coef.	Std. Err.	$p > \|z\|$	(95% Conf. Interval)	
Collision type	Collision with a guardrail	0.351	0.161	0.029	0.035	0.666
	Collision with other objects	0.835	0.283	0.003	0.281	1.389
Road surface	Wet-skid	0.207	0.137	0.031	−0.062	0.475
Lighting condition	Night without light	0.309	0.125	0.013	0.064	0.554
Terrain type	Mountainous	−0.695	0.321	0.030	−1.324	−0.066
Driver gender	Female	0.407	0.146	0.005	0.121	0.693
Driver age	55+ years	0.947	0.198	0.000	0.559	1.335
Driving experience	2- years	−1.094	0.202	0.000	−1.489	−0.699
	10+ years	−0.794	0.150	0.000	−1.087	−0.500
Vehicle type	Large vehicle	−0.885	0.167	0.000	−1.213	−0.557
Movement of vehicle	Not going straight	−1.570	0.405	0.000	−2.363	−0.777
cut1		−0.757	0.114		−0.981	−0.534
cut2		1.230	0.118		0.998	1.463

The absolute value of the coefficient can determine the order of effect of independent variables. The first three most influential factors were "vehicle not going straight", "driving experience 2- years", and "driver aged 55+ years". The three factors with minor influence were "wet-skid road surface", "night without light condition", and "vehicle collision with a guardrail".

The Brant test results in Table 4 show the p-value of five independent variables are less than 0.1. This indicates that the independent variable is not significant at the 90% confidence level; that is, the five independent variables violate the PLA. This proves the limitations of adopting the OL model and confirms the necessity of adopting the PPO model instead of a GOL model.

Only some of the independent variables violated PLA, so the GOL model could not be used directly here, and PPO played a key role as an intermediate coordination model. In order to reflect the effect of the improved model, the GOL model was also used here in addition to the OL and PPO models for analysis. The goodness of fit of the PPO model was compared with the OL and GOL model specifications using the AIC, BIC, and the LR test, as presented in Table 5. Comparing the PPO model to the OL and GOL models, there was a substantial difference (LR χ^2 = 299.39, p = 0.000). Due to its smallest AIC value and five additional parameters, the PPO model was chosen since it significantly outperformed the OL model. Additionally, the estimated full model performed better than the intercept-only model thanks to the predictor variables, which indicates that the results of parameter estimation were valid and the goodness of fit was appropriate (Pseudo R^2 = 0.1205).

Table 4. Results of Brant test of PLA.

Variable	Description	Chi2	p > chi2	Df
All		82.80	0.000	11
Road factors				
Road surface	Wet-skid	2.44	0.118	1
Terrain type	Mountainous	6.35	0.012	1
Environmental conditions				
Lighting condition	Night without light	0.47	0.493	1
Driver factors				
Driver gender	Female	17.24	0.000	1
Driver age	55+ years	0.01	0.920	1
Driving experience	2- years	7.61	0.006	1
	10+ years	2.63	0.105	1
Vehicle factors				
Collision type	Collision with a guardrail	3.40	0.115	1
	Collision with other objects	1.79	0.181	1
Vehicle type	Large vehicle	6.46	0.011	1
Movement of vehicle	Not going straight	17.69	0.000	1

Table 5. Results of the goodness-of-fit test.

Model	Log Likelihood	Df-AIC	LR	*AIC*	*BIC*	LR χ^2	Pseudo R^2
OL	−1131.356	13	5	2288.711	2354.374	221.09	0.0890
GOL	−1087.471	24	6	2222.941	2344.165	308.86	0.1244
PPO	−1092.208	18		2220.416	2311.334	299.39	0.1205

The parameters of the PPO model were optimized by the maximum likelihood method. The significance level for selecting explanatory variables was 0.10, and a stepwise selection process decided the significant explanatory variables. During the modeling process, variables were retained in the specification if they had t-statistics corresponding to the 90% confidence level or higher on a two-tailed t-test. Table 6 presents the parameter estimation results and marginal effects of all statistically significant variables on the probability of each severity level. The coefficients for nine and five predictor variables were statistically significant, respectively. A variable that violated the PLA had one beta coefficient, one gamma coefficient, and two alpha coefficients reflecting the cutoff points. Here, we focused on the statistically significant predictor factors.

Table 6. Results of factors affecting freeway crash injury severity based on the PPO model (n = 1154).

Variables		Coefficient		Marginal Effects (%)		
		β_j	γ_j	PDO	Injury	Fatal
Collision type	Collision with a guardrail	0.349 **		−6.46	0.64	5.82
	Collision with other objects	0.899 ***		−16.63	1.66	14.97
Road surface	Wet-skid	0.203		−3.76	0.37	3.38
Lighting condition	Night without light	0.314 **		−5.80	0.58	5.22
Terrain type	Mountainous	0.076	−2.896 ***	−1.40	49.60	−48.20
Driver gender	Female	1.148 ***	0.040 **	−21.23	20.56	0.67
Driver age	55+ years	1.008 ***		−18.64	1.86	16.78
Driving experience	2- years	−1.289 ***	−0.691 ***	23.83	−12.33	−11.51
	10+ years	−0.813 ***		15.03	−1.50	−13.53
Vehicle type	Large vehicle	−1.026 ***	−0.246 **	18.98	−14.88	−4.09
Movement of vehicle	Not going straight	−2.012 ***	−0.538 *	37.20	−28.25	−8.95
Alpha		0.763 ***	−1.216 ***			

*, **, ***indicate significance at 10%, 5%, and 1% levels, respectively.

5.2. Discussion

5.2.1. Effects of Driver Factors

The "female" variable violated the PLA. The gamma and the beta coefficients were 0.040 and 1.148, respectively. The coefficient yielded by adding these two numbers indicated that female drivers were more likely than male drivers to suffer severe injuries in crashes. A change from a male to a female driver raised the likelihood that the injury-severity level would fall into the injury category by 20.56%, and into the fatal category by 0.67%, according to the marginal effects for "female". Perhaps due to their innate psychology and personality features, female drivers are less able to withstand physical and emotional damage than male drivers [45,46]. Consistent conclusions have also been reached in previous studies [47]. Female drivers' mental capacity, emergency handling ability in the face of crisis situations, and resilience are lower than those of males, and their judgment of distance and speed is not as good as that of males. Male drivers are calmer in emergencies than female drivers are, and they also drive more flexibly and technically, which reduces the likelihood of serious injury. However, the frequency and probability of accidents involving no injuries will increase due to male drivers' proven capabilities, high-speed psychological needs, more aggressive driving style, and excessive confidence in their driving techniques.

The age of drivers greatly influenced the severity of the accident. The estimation results indicate that drivers aged 55+ years had a higher risk of severe injuries than drivers aged 26–40. A similar pattern was found in previous studies [47]. According to the marginal effects, there was a 1.86% rise in the likelihood of injury and a 16.78% increase in the likelihood of fatal when drivers aged 26–40 years were replaced by those aged 55+ years. As indicated in earlier research [17], older drivers who are more experienced and confident in driving and handling crises may, to some extent, cause accidents by carelessness or aggressive driving. There are no unanimously accepted conclusions from studies on the effect of driver age on accident severity [48,49]. The older drivers tend to react slowly to hazardous situations, may not be able to withstand crash impact forces well, may have a cognitive impairment, and may have other medical conditions. All or some of these factors may be responsible for their higher risk of injury severity. On the other hand, older drivers are more skilled, have more safety awareness, and are more patient when driving, which increases their likelihood of avoiding accidents involving no injuries. However, the gradual weakening of the physical function of the elderly and their slow response to emergencies lead to them having a higher probability of fatal accidents [17]. According to the age difference of those in accidents and preferences in life and entertainment, diversified traffic safety travel advertising or traffic accident emergency behavior instruction can be carried out.

The "driving experience 2- years" variable violated the PLA. The gamma and beta coefficients were −1.289 and −0.691. These two values were combined to yield the "driving experience 2- years" coefficient. The marginal effects indicated that for drivers with 3–10 years' experience and those with 2- years' experience the probability of injury or a fatality was reduced by 12.33 and 11.51%, respectively. Previous studies have found that drivers with 10- years of driving experience are more likely to be injured or killed in an accident [3]. The conclusion of this study digs further into the above findings and finds that drivers with very little driving experience, such as novice drivers with 2- years' experience, are instead relatively safer. This is consistent with the results of [47]. The decreased risk of injury among novice drivers may be due to their greater caution and lack of expertise on freeways. They drive much more slowly overall and strictly abide by the speed limit [49]. Of course, some other explanations are also possible. The efficiency of "driving experience 10+ years" indicates a decreased likelihood of sustaining severe injuries. The marginal effects show that a change from driving experience 3–10 years to driving experience 10+ years will reduce the probability of injury or fatality by 1.50 and 13.53%, respectively. This is consistent with the results of [3]. The potential reason is that drivers with driving experience 10+ years have more standardized driving behavior, shorter risk identification and decision-making times, and a lower probability of serious accidents.

In short, drivers with driving experience 3–10 years generally sustain more severe injury than other drivers because of their overconfidence in driving skills and risk perception. Training should also focus on psychological quality training, emergency response measures training, and enhanced driving standardization for drivers with less vehicle driving experience.

5.2.2. Effects of Vehicle Factors

Compared to collisions with vehicles, the coefficient for "collision with a guardrail" showed a higher risk of serious injuries. More specifically, the marginal effects indicate that a change from a collision between vehicles to a collision with a guardrail increased the probability that the injury-severity level was in the fatal category by 5.82%. Moreover, the coefficient of "collision with other objects" implied an increased likelihood of sustaining severe injuries in case of a collision with other objects compared to a collision with vehicles. The marginal effects indicate that a change from a collision between vehicles to a collision with other objects increased the probability that the injury-severity level was in the fatal category by 14.97%. When the driver reacts in an emergency, the vehicle may turn sharply, increasing the risk of a continuing accident and the likelihood of casualties. Compared to a collision with a guardrail, the likelihood of injury and fatality increased when colliding with other objects, reflecting the adverse effects of rigid obstacles such as piers on traffic safety.

Large vehicle crashes showed a lower likelihood of causing severe injuries. The marginal effects indicate that change from a small vehicle to a large vehicle being responsible reduced the probability of injury by 14.88%. Similar results were reported by previous studies [50,51]. The potential reason is that the driving speed of large vehicles is generally low, and the cab of large vehicles is usually high, which can protect the driver to a certain extent, thereby reducing the degree of accident damage. However, this contradicts the findings of some studies [20,52], and the results need further verification and research.

In terms of the vehicle's movement, those in a vehicle not going straight in a crash were at a lower risk of sustaining severe injury than those in one going straight. Finally, the marginal effects showed that a change in the movement of a vehicle from going straight to not going straight reduced the probability of injury by 28.25% and that of fatality by 8.95%. This result was expected because vehicles going straight will likely travel at higher speeds, and also the driver is always in a highly alert state when not going straight. Drivers are more reasonable in their handling of accidents in this situation, thus reducing their severity.

In taking speed limit measures, the traffic management department should set the appropriate speed limit value and speed limit mode in accordance with the running time and characteristics of large vehicles, and with prominent speed limit warning signs. It is recommended to install auxiliary stabilization devices for large vehicles to strengthen the stability of vehicle operation. At the same time, the detection of driver fatigue should be strengthened. If necessary, through a variety of communication channels, such as roadside signage, vehicle radio, etc., messages should improve driver awareness of safety precautions, as well as increase the seat belt wearing rate, because use of a seat belt in the event of an accident can also to a certain degree reduce the severity of the accident.

5.2.3. Effects of Road Factors

When a collision occurs on a wet-skid road surface compared to a dry road surface, the coefficient of "wet-skid" implies an increased probability of sustaining severe injuries. According to the marginal effects, switching from a dry road surface to one that was wet and slippery raised the likelihood that an injury would be fatal by 3.38%. This is consistent with previous studies that found abnormal road conditions can cause more serious accidents [50]. The likelihood of injury accidents increased because, even when the vehicle speed decreased in wet-skid conditions, the vehicle braking distance increased.

The "mountainous" variable was found to be violating the PLA. The gamma and beta coefficients were -2.896 and 0.076, respectively. These two were combined to yield the "mountainous" coefficient. According to marginal effects, the risk of the fatal injury category decreased by 48.20% when switching from a flat to a "mountainous" landscape. It was closely related to road sections' functional differences, traffic complexity, and structural diversity. More traffic interference will result from an uneven road layout, which increases the risk of injury or PDO accidents. Meanwhile, there is an increasing likelihood of fatal collisions on flat sections due to the faster speeds. Similar conclusions have been reached in previous studies [16,53]. Traffic managers should consider the effects of various road alignment characteristics and available locations when creating control measures. It is also required to improve safety levels, tighten traffic oversight and penalties, and optimize alignment design.

5.2.4. Effects of Environmental Conditions

Regarding light conditions, the coefficient of "night without light" indicated an increased likelihood of suffering severe injuries in case of a collision that happened at night without a light environment compared to a daylight environment. The worse the brightness was, the greater the impact on the accident's severity. Especially in the daytime, the probability of PDO accidents decreased by 5.80% due to a decrease in traffic volume, and the decrease of conflict points under the condition of no lighting at night. However, the probability of injury and fatal accidents increased by 0.58% and 5.22%, respectively. A similar finding was obtained in previous studies [6,47].

Poor traffic visibility and erratic traffic flow suggest that when there is insufficient lighting on the road, accidents are more likely to be serious ones [47,53]. Maintaining the required sight distance for safe freeway driving is difficult when there is no light at night, and the driver only relies on vehicle lights and reflective signs. The probability of rear-end collision or side collision increases without timely detection of the presence of vehicles in front [40]. The likelihood of fatigue in driving and inattention under dark conditions increases significantly. The accident's severity increases at night since the driver's reaction time increases, and the vehicle is more likely to travel at a higher speed. It is suggested to reasonably control the lighting time and light intensity of special road sections, increase the use of intelligent devices for light perception and sound control, repair faulty lamps in time, or add new light-sensitive materials to the relevant signs and markings. When visibility is low, the reflective level of the road signage can be appropriately strengthened through technical means, and the reflective level of the vehicle rear reflective signs can be enhanced, while reducing the vehicle speed limit.

6. Conclusions

This study aimed to analyze the influence of driver, vehicle, road, and environmental factors on crash injury severity on the freeway based on three different discrete choice models. Based on the data from this study, the fit of the PPO model was compared with the fit of the OL and GOL models. The variables that violated PLA were identified through the Brandt test. The PPO model was finally applied for detailed analysis because of its potency in handling a mixture of variables that met or violated the PLA. The results showed that collision with a guardrail and other objects, female drivers, and drivers aged 55+ years were more likely to cause injury and fatal outcomes. PDO was more likely, but injury and fatal accidents were less likely when the driving experience was 2- years, a large vehicle was responsible, and the vehicle was not going straight. Wet-skid road conditions enhanced the risk of injury accidents. The severity of an accident typically increased when it occurred at night on a road without lighting.

This study provides a new methodological reference for the field of freeway traffic safety research. The research results can provide a decision basis for the freeway management to take corresponding safety management measures in reducing the severity of freeway crashes. When rigid obstacles such as piers exist in the middle zone or lateral clearance range of a high-speed freeway, safety protection or energy absorption facilities should be added to reduce the severity of accidents. Dynamic variable speed limit signs and warnings can be installed to limit speed, improve driver vigilance, and reduce the accident severity when visibility declines or on wet-skid road surfaces. The lighting conditions of the corresponding sections should be improved to mitigate severe accidents in the accident-prone areas without lighting. When selecting the type of intermediate belt guardrail, a semi-rigid guardrail should be chosen as far as possible to meet the requirements of protection grade. Passive safety protection facilities can be added to the accident-prone sections with concrete guardrails. Strengthening drivers' safety education and training is necessary, especially for female drivers who have 3–10 years' experience. For senior drivers, it is vital to improve the assessment of their physical health status. Increased traffic enforcement and a decrease in lane-change violations by vehicles are required.

The research has some shortcomings. A variety of factors influence the likelihood of injury severity outcomes. This study only analyzed the factors for which data could be collected, but there are still many factors for which data are not easy to assemble or that have a significant impact on accidents but are unknown, such as instantaneous vehicle speed, real-time weather, use of seat belts, light conditions, etc. In addition, some of the influencing factors about human behavior were not considered in this study, such as fatigue, low performance, and driver vigilance and aggression, but the impact of these factors on the severity of accidents is very important. In follow-up study, more types of data can be collected for research to obtain more valuable research results. In addition to the PPO model, more models, such as the nested logit and RPL models, can be used to analyze accident severity. A comprehensive comparative analysis of the practicability of each model is also a direction for subsequent research. This study took only the traffic accidents on a particular freeway as a sample. It is fair to say there is a possibility that a wider area will provide insight into what has been documented in the present study. Different sample sizes will affect the parameter estimation of the model. This paper does not group the samples in-depth, and statistics of the sample sizes applicable to other models may also lead to deviations in the calibration results. In follow-up study, we can study the influence of different sample sizes on parameter estimation and group the samples to obtain better-fitting results.

Author Contributions: Visualization, W.C.; investigation, W.C.; writing—reviewing and editing, W.C.; data curation, F.Y.; methodology, F.Y.; software, F.Y.; formal analysis, C.W.; writing—original draft preparation, C.W.; validation, J.B.; software, J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China, grant number 2019YFE0108000, the General Research Project of the Zhejiang Provincial Department of Education, grant number Y202147371, and the National Social Science Foundation of China, grant number 19CGL004.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. Global Status Report on Road Safety 2015. *Inj. Prev.* **2015**, *15*, 286. [CrossRef]
2. Mannering, F.; Bhat, C. Analytic methods in accident research: Methodological frontier and future directions. *Anal. Methods Accid. Res.* **2014**, *1*, 1–22. [CrossRef]
3. Zhao, J.; Deng, W. Traffic accidents on expressways: New threat to China. *J. Crash Prev. Inj. Control* **2012**, *13*, 230–238. [CrossRef] [PubMed]
4. Wang, X.; Yang, S.; Yan, Y.; Li, X.; Du, J.; Tang, J. Analysis on distribution of freeway accidents under various conditions in China. *Adv. Mech. Eng.* **2016**, *8*, 241. [CrossRef]
5. Zeng, Q.; Sun, J.; Wen, H. Bayesian Hierarchical Modeling Monthly Crash Counts on Freeway Segments with Temporal Correlation. *J. Adv. Transp.* **2017**, *2017*, e5391054. [CrossRef]
6. Ye, F.; Cheng, W.; Wang, C.; Liu, H.; Bai, J. Investigating the severity of expressway crash based on the random parameter logit model accounting for unobserved heterogeneity. *Adv. Mech. Eng.* **2021**, *13*, 7278. [CrossRef]

7. Yu, R.; Abdel-Aty, M. Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Saf. Sci.* **2014**, *63*, 50–56. [CrossRef]
8. Ahmed, M.; Huang, H.; Abdel-Aty, M.; Guevara, B. Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway. *Accid. Anal. Prev.* **2011**, *43*, 1581–1589. [CrossRef]
9. Ma, X.; Chen, F.; Chen, S. Modeling Crash Rates for a Mountainous Highway by Using Refined-Scale Panel Data. *Transp. Res. Rec.* **2015**, *2515*, 10–16. [CrossRef]
10. Wen, H.; Sun, J.; Zeng, Q.; Zhang, X.; Yuan, Q. The Effects of Traffic Composition on Freeway Crash Frequency by Injury Severity: A Bayesian Multivariate Spatial Modeling Approach. *J. Adv. Transp.* **2018**, *2018*, e6964828. [CrossRef]
11. Hou, Q.; Tarko, A.P.; Meng, X. Analyzing crash frequency in freeway tunnels: A correlated random parameters approach. *Accid. Anal. Prev.* **2018**, *111*, 94–100. [CrossRef] [PubMed]
12. Ye, X.; Pendyala, R.M.; Shankar, V.; Konduri, K.C. A simultaneous equations model of crash frequency by severity level for freeway sections. *Accid. Anal. Prev.* **2013**, *57*, 140–149. [CrossRef]
13. Hou, Q.; Meng, X.; Leng, J.; Yu, L. Application of a random effects negative binomial model to examine crash frequency for freeways in China. *Phys. A* **2018**, *509*, 937–944. [CrossRef]
14. Ma, Z.; Zhang, H.; Chien, S.I.J.; Wang, J.; Dong, C. Predicting expressway crash frequency using a random effect negative binomial model: A case study in China. *Accid. Anal. Prev.* **2017**, *98*, 214–222. [CrossRef]
15. Hou, Q.; Tarko, A.P.; Meng, X. Investigating factors of crash frequency with random effects and random parameters models: New insights from Chinese freeway study. *Accid. Anal. Prev.* **2018**, *120*, 1–12. [CrossRef]
16. Osman, M.; Mishra, S.; Paleti, R. Injury severity analysis of commercially-licensed drivers in single-vehicle crashes: Accounting for unobserved heterogeneity and age group differences. *Accid. Anal. Prev.* **2018**, *118*, 289–300. [CrossRef]
17. Kim, J.K.; Ulfarsson, G.F.; Kim, S.; Shankar, V.N. Driver-injury severity in single-vehicle crashes in California: A mixed logit analysis of heterogeneity due to age and gender. *Accid. Anal. Prev.* **2013**, *50*, 1073–1081. [CrossRef]
18. Radzi, N.H.M.; Gwari, I.S.B.; Mustaffa, N.H.; Sallehuddin, R. Support Vector Machine with Principle Component Analysis for Road Traffic Crash Severity Classification. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *551*, 012068. [CrossRef]
19. Se, C.; Champahom, T.; Jomnonkwao, S.; Ratanavaraha, V. Motorcyclist injury severity analysis: A comparison of Artificial Neural Networks and random parameter model with heterogeneity in means and variances. *Int. J. Inj. Control Saf. Promot.* **2022**, 1–16. [CrossRef] [PubMed]
20. Zeng, Q.; Gu, W.; Zhang, X.; Wen, H.; Lee, J.; Hao, W. Analyzing freeway crash severity using a Bayesian spatial generalized ordered logit model with conditional autoregressive priors. *Accid. Anal. Prev.* **2019**, *127*, 87–95. [CrossRef] [PubMed]
21. Chu, H.C. Assessing factors causing severe injuries in crashes of high-deck buses in long-distance driving on freeways. *Accid. Anal. Prev.* **2014**, *62*, 130–136. [CrossRef] [PubMed]
22. McCullagh, P. *Generalized Linear Models*, 2nd ed.; Routledge: New York, NY, USA, 1983. [CrossRef]
23. Peterson, B.; Harrell, F.E., Jr. Partial Proportional Odds Models for Ordinal Response Variables. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1990**, *39*, 205–217. [CrossRef]
24. Falcone, D.; Bona, G.D.; Duraccio, V.; Silvestri, A. Integrated hazards method (IHM): A new safety allocation technique. In Proceedings of the Iasted International Conference on Modelling & Simulation, Montreal, QC, Canada, 30 May–1 June 2007; Available online: http://www.researchgate.net/publication/262210653_Integrated_hazards_method_IHM_a_new_safety_allocation_technique (accessed on 6 January 2023).
25. Forcina, A.; Silvestri, L.; Di Bona, G.; Silvestri, A. Reliability allocation methods: A systematic literature review. *Qual. Reliab. Eng. Int.* **2020**, *36*, 2085–2107. [CrossRef]
26. Chen, Z.; Fan, W.D. A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. *Int. J. Transp. Sci. Technol.* **2019**, *8*, 43–52. [CrossRef]
27. Ratanavaraha, V.; Suangka, S. Impacts of accident severity factors and loss values of crashes on expressways in Thailand. *IATSS Res.* **2014**, *37*, 130–136. [CrossRef]
28. Mergia, W.; Eustace, D.; Chimba, D.; Qumsiyeh, M. Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio. *Accid. Anal. Prev.* **2013**, *55*, 202–210. [CrossRef]
29. Savolainen, P.T.; Mannering, F.L.; Lord, D.; Quddus, M.A. The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accid. Anal. Prev.* **2011**, *43*, 1666–1676. [CrossRef]
30. Zhou, Z.; Meng, F.; Song, C.; Richard, T.; Guo, Z.; Yang, L.; Wang, W. Factors associated with consecutive and non-consecutive crashes on freeways: A two-level logistic modeling approach. *Accid. Anal. Prev.* **2021**, *154*, 106054. [CrossRef]
31. Liu, P.; Fan, W. Modeling head-on crash severity on NCDOT freeways: A mixed logit model approach. *Can. J. Civ. Eng.* **2019**, *46*, 322–328. [CrossRef]
32. Haleem, K.; Gan, A. Effect of driver's age and side of impact on crash severity along urban freeways: A mixed logit approach. *J. Saf. Res.* **2013**, *46*, 67–76. [CrossRef]
33. Yan, L.; He, Y.; Qin, L.; Wu, C.; Zhu, D.; Ran, B. A novel feature extraction model for traffic injury severity and its application to Fatality Analysis Reporting System data analysis. *Sci. Prog.* **2020**, *103*, 6471. [CrossRef] [PubMed]
34. Wu, Q.; Zhang, G.; Zhu, X.; Liu, X.C.; Tarefder, R. Analysis of driver injury severity in single-vehicle crashes on rural and urban roadways. *Accid. Anal. Prev.* **2016**, *94*, 35–45. [CrossRef] [PubMed]

35. Adanu, E.K.; Hainen, A.; Jones, S. Latent class analysis of factors that influence weekday and weekend single-vehicle crash severities. *Accid. Anal. Prev.* **2018**, *113*, 187–192. [CrossRef] [PubMed]
36. Ma, Z.; Chien, S.I.J.; Dong, C.; Hu, D.; Xu, T. Exploring factors affecting injury severity of crashes in freeway tunnels. *Tunn. Undergr. Space Technol.* **2016**, *59*, 100–104. [CrossRef]
37. Wang, Z.; Chen, H.; Lu, J.J. Exploring impacts of factors contributing to injury severity at freeway diverge areas. *Transp. Res. Rec.* **2009**, *2102*, 43–52. [CrossRef]
38. Ye, F.; Wang, C.; Cheng, W.; Liu, H. Exploring factors associated with cyclist injury severity in vehicle-electric bicycle crashes based on a random parameter logit model. *J. Adv. Transp.* **2021**, *2021*, e5563704. [CrossRef]
39. Ma, Z.; Zhao, W.; Chien, S.I.J.; Dong, C. Exploring factors contributing to crash injury severity on rural two-lane highways. *J. Saf. Res.* **2015**, *55*, 171–176. [CrossRef]
40. Pervez, A.; Lee, J.; Huang, H. Exploring factors affecting the injury severity of freeway tunnel crashes: A random parameters approach with heterogeneity in means and variances. *Accid. Anal. Prev.* **2022**, *178*, 106835. [CrossRef] [PubMed]
41. Angiulli, F.; Fassetti, F. Distance-based outlier queries in data streams: The novel task and algorithms. *Data Min. Knowl. Disc.* **2010**, *20*, 290–324. [CrossRef]
42. McCullagh, P. Regression Models for Ordinal Data. *J. R. Stat. Soc. Ser. B (Methodol.)* **1980**, *42*, 109–127. [CrossRef]
43. Williams, R. Understanding and interpreting generalized ordered logit models. *J. Math. Sociol.* **2016**, *40*, 7–20. [CrossRef]
44. Liu, X.; Koirala, H. Ordinal regression analysis: Using generalized ordinal logistic regression models to estimate educational data. *J. Mod. Appl. Stat. Methods* **2012**, *11*, 242–254. [CrossRef]
45. Chen, H.; Cao, L.; Logan, D.B. Analysis of Risk Factors Affecting the Severity of Intersection Crashes by Logistic Regression. *J. Crash Prev. Inj. Control* **2012**, *13*, 300–307. [CrossRef] [PubMed]
46. Sivak, M.; Schoettle, B.; Rupp, J. Survival in Fatal Road Crashes: Body Mass Index, Gender, and Safety Belt Use. *J. Crash Prev. Inj. Control* **2010**, *11*, 66–68. [CrossRef] [PubMed]
47. Hou, Q.; Huo, X.; Leng, J.; Cheng, Y. Examination of driver injury severity in freeway single-vehicle crashes using a mixed logit model with heterogeneity-in-means. *Phys. A* **2019**, *531*, 121760. [CrossRef]
48. Doroudgar, S.; Chuang, H.M.; Perry, P.J.; Thomas, K.; Bohnert, K.; Canedo, J. Driving performance comparing older versus younger drivers. *Traffic Inj. Prev.* **2017**, *18*, 41–46. [CrossRef]
49. Moták, L.; Gabaude, C.; Bougeant, J.C.; Huet, N. Comparison of driving avoidance and self-regulatory patterns in younger and older drivers. *Transp. Res. F Traffic Psychol. Behav.* **2014**, *26*, 18–27. [CrossRef]
50. Yamamoto, T.; Shankar, V.N. Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. *Accid. Anal. Prev.* **2004**, *36*, 869–876. [CrossRef]
51. Shao, X.; Ma, X.; Chen, Q.; Song, S.; Pan, L.; You, K. A random parameters ordered probit analysis of injury severity in truck involved rear-end collisions. *Int. J. Environ. Res. Public Health* **2020**, *17*, 395. [CrossRef]
52. Zeng, Q.; Wen, H.; Huang, H. The interactive effect on injury severity of driver-vehicle units in two-vehicle crashes. *J. Saf. Res.* **2016**, *59*, 105–111. [CrossRef]
53. Christoforou, Z.; Cohen, S.; Karlaftis, M.G. Vehicle occupant injury severity on highways: An empirical investigation. *Accid. Anal. Prev.* **2010**, *42*, 1606–1620. [CrossRef] [PubMed]