sustainability

MDPI

Article

Modeling the Unobserved Heterogeneity in E-bike Collision Severity Using Full Bayesian Random Parameters Multinomial Logit Regression

Yanyong Guo [1,2,*], Yao Wu [1], Jian Lu [1] and Jibiao Zhou [3]

1. Jiangsu Key Laboratory of Urban ITS, Jiangsu Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University, Si Pai Lou No. 2, Nanjing 210096, China; wuyaomaster@126.com (Y.W.); lujian_1972@seu.edu.cn (J.L.)
2. Department of Civil Engineering, The University of British Columbia, 6250 Applied Science Lane, Vancouver, BC V6T 1Z4, Canada
3. School of Civil and Transportation Engineering, Ningbo University of Technology, Ningbo 315211, China; zhoujb2014@nbut.edu.cn

* Correspondence: yanyong.guo@ubc.ca

Received: 17 March 2019; Accepted: 2 April 2019; Published: 8 April 2019

check for updates

Abstract: Understanding the risk factors of e-bike collisions can improve e-bike riders' safety awareness and help traffic professionals to develop effective countermeasures. This study investigates risk factors that significantly contribute to the severity of e-bike collisions. Two months of e-bike collision data were collected in the city of Ningbo, China. A random parameters multinomial logit regression (RP-MNL) is proposed to account for the unobserved heterogeneity across observations. A fixed parameters multinomial logit regression (FP-MNL) is estimated and compared with the RP-MNL under the Bayesian framework. The full Bayesian approach based on Markov chain Monte Carlo simulation is employed to estimate the model parameters. Both parameter estimates and odds ratio (OR) are used to interpret the impact of risk factors on the severity of e-bike collisions. The model comparison results show that RP-MNL outperforms FP-MNL, indicating that accommodating the unobserved heterogeneity across observations could improve the model fit. The model estimation results show that age, gender, e-bike behavior, license plate, bicycle type, location, and speed limit are statistically significant and associated with the severity of e-bike collisions. Furthermore, four risk factors, i.e., gender, e-bike behavior, bicycle type, and speed limit, are found to have heterogeneous effects on severity of e-bike collisions, appearing in the form of random parameters in the statistical model.

Keywords: E-bike collision; severity; unobserved heterogeneity; random parameters multinomial logit model; full Bayesian approach

1. Introduction

E-bikes are becoming a rising transportation mode for commuting in China. E-bike ownership increased significantly from 58 thousand in 1988 to 466 million in 2016, with an annual increase of 64.8% [1]. Due to their small size and flexible routes, e-bikes provide road users with convenient and affordable mobility. Furthermore, they are beneficial to the public for the advantages of low cost and easy parking [2]. However, a big issue preventing the use of e-bikes is the increasing number of e-bike collisions [3]. Particularly, e-bikes can result in serious collisions due to their lack of protection. As such, it is important to understand risk factors related to e-bike collisions in order to increase e-bike riders' safety awareness and help traffic professionals develop effective countermeasures.

A number of risk factors such as traffic and road/intersection design factors are known to affect e-bike collisions [4–10]. However, there is unobserved heterogeneity across observations, which may affect severity of e-bike collisions. Take gender as an example. A traditional approach considers gender as an observed human element affecting e-bike collisions. However, even within the same gender group, there is great variation because of the difference in height, weight and other physiological factors [11]. If such unobserved heterogeneity is ignored, the modeling results of e-bike collisions can be inaccurate and the statistical model may result in biased and erroneous parameters estimates.

The objective of the study is to investigate risk factors that significantly contribute to severity of e-bike collisions. A random parameters multinomial logit regression (RP-MNL) was developed to account for the unobserved heterogeneity across observations. The full Bayesian method based on Markov chain Monte Carlo (MCMC) simulation was utilized for the estimation of RP-MNL. For the purpose of comparison, a fixed parameters multinomial logit regression (FP-MNL) was estimated and compared with the RP-MNL under the Bayesian framework.

2. Literature Review

Due to the increased use of e-bikes and their high fatality rate, e-bike safety has been increasingly recognized by researchers. Weinert et al. [12] examined the e-bike riders' safety perceptions in the city of Shijiazhuang, China. The findings showed that female riders feel safer riding e-bikes to cross intersections than riding bicycles. Feng et al. [3] investigated the change of e-bike injuries in China. The results showed that e-bike injuries and deaths were increasing, while traditional bicycle injuries and deaths were decreasing.

Yao and Wu [4] established the correlations among safety attitude, risk perception, and aberrant riding behaviors of e-bike riders using the data from a self-reported questionnaire survey. The findings suggested that males had a higher likelihood of being involved in at-fault collisions than females. Moreover, e-bike riders who have a driver license were less likely to be involved in crashes. Du et al. [8] investigated e-bike riders' illegal behaviors in Suzhou, China. The results showed that e-bike riders were used to violating traffic rules, and this behavior can lead to a high crash risk.

Lawinger and Bastian [5] examined e-bike crashes in Germany and concluded that e-bike collisions were more severe than bicycle collisions. However, another study conducted by Otte et al. [13] showed that there were no significant differences in injury propensity and injury severity between e-bike collisions and bicycle collisions in Germany. Weber et al. [7] examined the impacts of risk factors, including type of accident, helmet usage, and injury severity, on e-bike collisions using the police-recorded accidents data in Switzerland. Papoutsi et al. [14] analyzed e-bike collisions by patients' age and gender, and cause of the accident using the hospital data in Switzerland. The results showed the majority of patients were male and the main causes of injury were self-accident.

Hu et al. [6] explored the risk factors associated with injury severity of e-bikes and bicycles in Hefei, China. The findings showed that age, gender, and vehicle type were related with e-bike and bicycle injuries. Schepers et al. [15] conducted a case-control study to compare the likelihood of collisions between e-bike riders and bicyclists. The results showed that e-bike riders were more likely to be involved in a collision than bicyclists. Guo et al. [16] analyzed e-bikes red-light running behaviors and found that male e-bike riders were more likely to have risky behaviors.

Wang et al. [9] modeled the faults among e-bike fatal collisions in China. The results showed that pre-crash behaviors of both drivers and e-bike riders were significantly related to fault assignment. Guo et al. [10] evaluated the factors affecting e-bikes involved crashes and license plate use in China. The findings showed that gender, age, education level, driver license, car in household, experiences in using e-bike, law compliance, and aggressive driving behaviors were significantly related to both e-bikes involved crashes and license plate use.

3. Date

E-bike collision data were collected from the city of Ningbo, China. E-bikes have been used as a common commuting transportation mode in Ningbo. E-bike ownership in the urban area reached 26.5 thousand in 2015. However, due to the high number of e-bike collisions, e-bike safety has become a public safety concern. In 2015, 1541 e-bike collisions were reported by the Ningbo Police Department (NBPD). The number of e-bike collisions was almost twice that of bicycle collisions (854) and was comparable to motorcycle collisions (1676). However, the fatality proportion of e-bike collisions (4.5%) is much higher than those of motorcycle collisions (2.3%) and bicycle collisions (3.8%).

E-bike collision data were provided by NBPD. The crash database includes rich information. Within each record, the data, time, location, collision type, crash severity, involved vehicle type, weather, lights condition, intentions and behaviors of involved drivers, driver's age, gender, ID number, license plate number and telephone number were available. Two months of collision data (July and August in 2015) were extracted from the crash database, which was used in a previous study [10]. To filter the e-bike collision data, at least one of the collision objects should be e-bike in the collision record.

Finally, a total of 310 e-bike collision records were extracted and used for the analysis. Among them, 7 were reported as fatality (F), accounting for 2.26% of the total e-bike collisions, 43 (13.87%) were reported as incapacitating injury (I), 233 (75.16%) were reported as non-incapacitating injury (NI), and 27 (8.71%) were reported as property damage only (PDO). Due to the small sample size of fatality collisions, they are combined with the incapacity injury. As such, the collision severity was reorganized into three categories: F/I, NI, and PDO.

In addition to the original information included in the crash database, the collision location (coordinates) was used to match the collision points on the Google Earth map. As such, the geometric design factors such as road type, median type, and width were collected. Table 1 shows detailed information regarding variable definitions and descriptions, as well as the percentage of observed collision frequency at each severity level.

Table 1. Descriptive statistics of e-bike collisions severity and explanatory variables.

Variable	Categories		No. of Collisions	Collisions Severity (%)		
	Value	Description		F/I	N	PDO
E-bike collisions			310	16.13	75.16	8.71
E-bike rider characteristics						
Age	0	Young	137	10.95	72.26	16.79
	1	Middle-aged	125	14.40	75.20	10.40
	2	Old *	48	16.67	77.08	6.25
Gender	0	Male	188	18.62	76.06	5.32
	1	Female *	122	14.75	73.77	11.48
E-bike behavior	0	Violation	117	17.09	76.07	6.84
	1	Distraction	67	17.91	74.63	7.46
	2	Normal riding *	126	11.11	70.63	18.25
License plate	0	Yes	114	15.79	74.56	9.65
	1	No *	196	16.33	77.04	6.63
Bicycle type	0	E-bike	168	17.26	73.21	9.52
	1	E-scooter *	142	14.79	77.46	7.75
Roadway characteristics						
Road type	0	Minor road	123	15.45	77.24	7.32
	1	Major road *	187	17.11	73.26	9.63
Location	0	Intersection	227	17.18	76.21	6.61
	1	Segment *	83	15.66	74.70	9.64
Median type	0	Divided	255	17.65	71.37	10.98
	1	Undivided *	55	12.73	76.36	10.91
Median width	1	≥1 m	212	18.40	72.64	8.96
	0	<1 m *	98	14.29	77.55	8.16
Surface type	0	Asphalt	198	17.17	76.77	6.06
	1	Other *	112	15.18	75.00	9.82
Speed limit	0	≥45 km/h	167	20.36	73.65	5.99
	1	<45 km/h *	143	12.59	77.62	9.79

* Selected as the base of the categorical variable.

4. Methodology

4.1. Multinomial Logit Regression

The multinomial logit regression (MNL) is commonly used in collision severity analysis, in which collisions can be categorized into more than two levels with one level as a reference category [17]. In this study, an MNL is developed to explore the risk factors with different severity levels of e-bike collisions, with the PDO e-bike collisions as reference level. The MNL is expressed as

$$P(Y_i = j) = \frac{\exp\left(\alpha_0^j + \alpha_1^j x_{i1} + \alpha_2^j x_{i2} + \cdots + \alpha_K^j x_{iK}\right)}{\sum\limits_{j=1}^{J} \exp\left(\alpha_0^j + \alpha_1^j x_{i1} + \alpha_2^j x_{i2} + \cdots + \alpha_K^j x_{iK}\right)} \tag{1}$$

where $Y_i = j$ is the collision severity j for the ith observation; $\boldsymbol{X} = [x_{i1}, x_{i2}, \ldots, x_{iK}]$ is the $1 \times K$ vector of explanatory variables; $\boldsymbol{\alpha} = [\alpha_0^j, \alpha_1^j, \alpha_2^j, \cdots, \alpha_K^j]^T$ is the coefficient vector for the vector $\boldsymbol{X}$.

The likelihood function for MNL is given as

$$f(\boldsymbol{Y}|\boldsymbol{\alpha}) = \prod_{i=1}^{N}\prod_{j=1}^{J}[\delta_{ij} \times P(Y_i = j)] = \prod_{i=1}^{N}\prod_{j=1}^{J}\left[\delta_{ij} \times \frac{\exp\left(\alpha_0^j + \alpha_1^j x_{i1} + \alpha_2^j x_{i2} + \cdots + \alpha_K^j x_{iK}\right)}{\sum\limits_{j=1}^{J} \exp\left(\alpha_0^j + \alpha_1^j x_{i1} + \alpha_2^j x_{i2} + \cdots + \alpha_K^j x_{iK}\right)}\right] \tag{2}$$

where N is the sample size; J is the total number of outcomes (e-bike collision severity levels); δ_{ij} is an indicator which equals to 1 if the discrete outcome for sample i is j, and 0 otherwise.

4.2. Random Parameters Multinomial Logit Regression

The parameters of explanatory variables in standard MNL are assumed to be fixed across observations, indicating that the impact of each explanatory variable is the same across observations [18]. However, this assumption is somehow contrary to the fact that the effect of explanatory variable varies across observations. To handle the unobserved heterogeneity across observations, this study proposes a random parameters multinomial logit regression (RP-MNL) to investigate the risk factors affecting e-bikes collision severity of e-bikes collisions. Compared with the fixed-parameters MNL (FP-MNL), the RP-MNL allows all parameters to vary randomly across observations. As such, more features in the collision data can be extracted and the accuracy of the regression can be improved. The RP-MNL is developed by being written Equation (1) as follows

$$P(Y_i = j) = \frac{\exp\left(\alpha_{i,0}^j + \alpha_{i,1}^j x_{i1} + \alpha_{i,2}^j x_{i2} + \cdots + \alpha_{i,K}^j x_{iK}\right)}{\sum\limits_{j=1}^{J} \exp\left(\alpha_{i,0}^j + \alpha_{i,1}^j x_{i1} + \alpha_{i,2}^j x_{i2} + \cdots + \alpha_{i,K}^j x_{iK}\right)} \tag{3}$$

where $\boldsymbol{\alpha} = [\alpha_{i,0}^j, \alpha_{i,1}^j, \alpha_{i,2}^j, \cdots, \alpha_{i,K}^j]^T$ is the coefficient vector, and these random parameters are allowed to vary across observations.

In this study, the random parameters in RP-MNL are assumed to be normally distributed as $\boldsymbol{\alpha} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} = [\mu_0^j, \mu_1^j, \mu_2^j, \cdots, \mu_K^j]^T$ and $\boldsymbol{\Sigma} = [\Sigma_0^j, \Sigma_1^j, \Sigma_2^j, \cdots, \Sigma_K^j]^T$. Specifically,

$$\begin{bmatrix} \alpha_{i,0}^{j} \\ \alpha_{i,1}^{j} \\ \alpha_{i,2}^{j} \\ \vdots \\ \alpha_{i,K}^{j} \end{bmatrix} \sim N \begin{bmatrix} \mu_0^j, \Sigma_0^j \\ \mu_1^j, \Sigma_1^j \\ \mu_2^j, \Sigma_2^j \\ \vdots \\ \mu_K^j, \Sigma_K^j \end{bmatrix} \tag{4}$$

Similarly, the likelihood of the RP-MNL is given as

$$f(\boldsymbol{Y}|\boldsymbol{\alpha}) = \prod_{i=1}^{N}\prod_{j=1}^{J}[\delta_{ij} \times P(Y_i = j)] = \prod_{i=1}^{N}\prod_{j=1}^{J}\left[\delta_{ij} \times \frac{\exp\left(\alpha_{i,0}^{j} + \alpha_{i,1}^{j}x_{i1} + \alpha_{i,2}^{j}x_{i2} + \cdots + \alpha_{i,K}^{j}x_{iK}\right)}{\sum\limits_{j=1}^{J}\exp\left(\alpha_{i,0}^{j} + \alpha_{i,1}^{j}x_{i1} + \alpha_{i,2}^{j}x_{i2} + \cdots + \alpha_{i,K}^{j}x_{iK}\right)}\right] \tag{5}$$

4.3. Full Bayesian Estimation

The full Bayesian approach based on Markov chain Monte Carlo (MCMC) is utilized to estimate the RP-MNL. In the full Bayesian approach, prior information and observed data are combined to obtain the RP-MNL parameters' posterior distributions. Let $\boldsymbol{\Theta}$ represent the parameters in RP-MNL, which is given as follows

$$\Theta = [\boldsymbol{\alpha}, \boldsymbol{\mu}, \sum] \tag{6}$$

According to the Bayesian inference, the posterior distribution of parameters $\boldsymbol{\Theta}$ can be estimated as follows

$$f(\boldsymbol{\Theta}|\boldsymbol{Y}) = \frac{f(\boldsymbol{Y},\boldsymbol{\Theta})}{f(\boldsymbol{Y})} = \frac{f(\boldsymbol{Y}|\boldsymbol{\Theta})\pi(\boldsymbol{\Theta})}{\int f(\boldsymbol{Y},\boldsymbol{\Theta})d\boldsymbol{\Theta}} \propto f(\boldsymbol{Y}|\boldsymbol{\Theta})\pi(\boldsymbol{\Theta}) \tag{7}$$

where $f(\boldsymbol{\Theta} \mid \boldsymbol{Y})$ is the posterior distribution of parameters $\boldsymbol{\Theta}$ conditional on observed dataset $\boldsymbol{Y}$; $f\,(\boldsymbol{Y}, \boldsymbol{\Theta})$ is the joint probability distribution of observed dataset $\boldsymbol{Y}$ and parameters $\boldsymbol{\Theta}$; $\pi(\boldsymbol{\Theta})$ is the prior distribution of parameters $\boldsymbol{\Theta}$; $f\,(\boldsymbol{Y} \mid \boldsymbol{\Theta})$ is the likelihood conditional function based on parameters $\boldsymbol{\Theta}$.

Due to the lack of information of the random parameters, the non-informative prior distributions for the random parameters in RP-MNL are used in this study. The prior distributions for parameters are given as follows

$$\mu_k^j \sim N(\overline{a}_k^j, \overline{b}_k^j) \tag{8}$$

$$\Sigma_k^j \sim Inverse\ \ gamma(\overline{c}_k^j, \overline{d}_k^j) \tag{9}$$

where the mean of the random parameters follows normal distribution, and the variance of the random parameters follows inverse gamma distribution.

The parameters with over lines in Equations (8) and (9) are hyper-parameters which are given as follows

$$\overline{a}_k^j = 0,\ \ \overline{b}_k^j = 10^6 \tag{10}$$

$$\overline{c}_k^j = 10^{-3},\ \ \overline{d}_k^j = 10^{-3} \tag{11}$$

Based on the prior distributions of parameters $\boldsymbol{\Theta}$, the posterior distribution $f(\boldsymbol{\Theta} \mid \boldsymbol{Y})$ can be derived as follows

$$\begin{aligned} & f(\boldsymbol{\Theta}|\boldsymbol{Y}) \propto f(\boldsymbol{Y}|\boldsymbol{\Theta})\pi(\boldsymbol{\Theta}) \propto f(\boldsymbol{Y}|\boldsymbol{\alpha})\pi(\boldsymbol{\Theta}) \\ & = f(\boldsymbol{Y}|\boldsymbol{\alpha}) \times \prod_{i=1}^{N}\prod_{k=0}^{K}\prod_{j=1}^{J} N(\alpha_{i,k}^{j}\Big|u_k^j, \Sigma_k^j) \times \prod_{k=0}^{K}\prod_{j=1}^{J} N(u_k^j\Big|\overline{a}_k^j, \overline{b}_k^j) \times \prod_{k=0}^{K}\prod_{j=1}^{J} IG(\Sigma_k^j\Big|\overline{c}_k^j, \overline{d}_k^j) \\ & = \prod_{i=1}^{N}\prod_{j=1}^{J}\left[\delta_{ij} \times \frac{\exp\left(\alpha_{i,0}^j + \alpha_{i,1}^j x_{i1} + \alpha_{i,2}^j x_{i2} + \cdots + \alpha_{i,K}^j x_{iK}\right)}{\sum\limits_{j=1}^{J}\exp\left(\alpha_{i,0}^j + \alpha_{i,1}^j x_{i1} + \alpha_{i,2}^j x_{i2} + \cdots + \alpha_{i,K}^j x_{iK}\right)}\right] \times \prod_{i=1}^{N}\prod_{k=0}^{K}\prod_{j=1}^{J}\left[\frac{1}{\sqrt{2\pi\Sigma_k^j}}\exp\left(-\frac{1}{2}\frac{\left(\alpha_k^j - u_k^j\right)^2}{\Sigma_k^j}\right)\right] \\ & \times \prod_{k=0}^{K}\prod_{j=1}^{J}\left[\frac{1}{\sqrt{2\pi 10^3}}\exp\left(-\frac{1}{2}\frac{\left(u_k^j\right)^2}{10^6}\right)\right] \times \prod_{k=0}^{K}\prod_{j=1}^{J}\frac{\Sigma_k^{j(10^{-3}-1)}\left(10^{-3}\right)^{10^{-3}}}{\Gamma(10^{-3})}\exp(-10^{-3} \times \Sigma_k^j) \end{aligned} \tag{12}$$

4.4. Risk Factors Analysis

The odds ratio (OR) is employed to analyze the impact of risk factors on the e-bike collision severities. The OR of a risk factor means an increase in the odds of the outcome (severity levels) if the value of the risk factor increases by one unit [19]. The OR for a risk factor x_j in the RP-MNL can be calculated as follows

$$\mathrm{OR} = \frac{odds\left(\boldsymbol{X}, x_j + 1\right)}{odds\left(\boldsymbol{X}, x_j\right)} = \frac{\exp(\boldsymbol{\alpha X}) \times \exp\left(\alpha_j\right)}{\exp(\boldsymbol{\alpha X})} = \exp\left(\alpha_j\right) \tag{13}$$

4.5. Models Comparison

The Deviance Information Criteria (DIC) is used to compare the full Bayesian estimated regressions. The DIC is given as follows

$$DIC = \overline{D} + p_D;\ p_D = \overline{D} - \hat{D} \tag{14}$$

where D represents the unstandardized deviance of the postulated model; $\overline{D}$ represents the posterior mean of D; $\hat{D}$ represents the point estimate obtained by substituting the posterior means of the model's parameters in D; p_D is a measure of model complexity. The model with a smaller DIC outperforms the model with a larger DIC.

5. Results and Discussion

5.1. Estimation Results

Both FP-MNL and RP-MNL were estimated to identify and evaluate the impact of risk factors on severity of e-bike collisions. The performance of these two models were compared using DIC. The MCMC simulation-based full Bayesian approach was employed to estimate the posterior distributions of the models' parameters. WinBUGS software was used as the modeling platform. Two independent Markov chains for each of the parameters with diverse initial values run for 40,000 iterations. The first 20,000 iterations in each chain were used for monitoring convergence and then discarded as burn-in runs. The convergence of the posterior distribution was monitored by visual inspection of the trace, and autocorrelation plots. In this study, the collision severity PDO was selected as the reference category. Tables 2 and 3 show the estimates results of the FP-MNL and the RP-MNL respectively. The variables which are significant at 95% credible interval were kept in the tables.

Table 2. Estimates of parameters in FP-MNL.

Variables	F/I					NI				
	Mean	Std.	95% Credible Interval		OR	Mean	Std.	95% Credible Interval		OR
			Lower	Upper				Lower	Upper	
Age										
Young vs. Old	0.87	0.079	0.71	1.14	2.35	0.44	0.057	0.3	0.55	1.56
Middle-aged vs. Old	0.34	0.057	0.21	0.51	1.44	0.36	0.048	0.26	0.47	1.42
Gender										
Male vs. Female	0.67	0.027	0.55	0.75	1.91	1.15	0.083	0.97	1.35	2.91
E-bike Behavior										
Violation vs. Normal riding	1.22	0.446	0.33	2.15	3.22	1.01	0.286	0.45	1.65	2.75
Distraction vs. Normal riding	0.94	0.328	0.29	1.58	2.45	0.79	0.227	0.33	2.16	2.01
License Plate or Not	−0.53	0.225	−0.97	−0.09	0.58	−0.51	0.238	−0.97	−0.04	0.61
Yes vs. No										
Bicycle Type	−0.43	0.127	−0.66	−0.18	0.67	−0.49	0.166	−0.82	−0.16	0.57
E-bike vs. E-scooter										
Location	0.75	0.258	0.24	1.25	2.14	0.66	0.272	0.13	1.21	1.78
Intersection vs. Segment										
Speed Limit	0.56	0.093	0.37	0.75	1.82	0.62	0.142	0.34	0.89	1.88
≥45 km/h vs. <45 km/h	0.87	0.079	0.71	1.14	2.35	0.44	0.057	0.3	0.55	1.56
DIC	457.3									

Table 3. Estimates of parameters in RP-MNL.

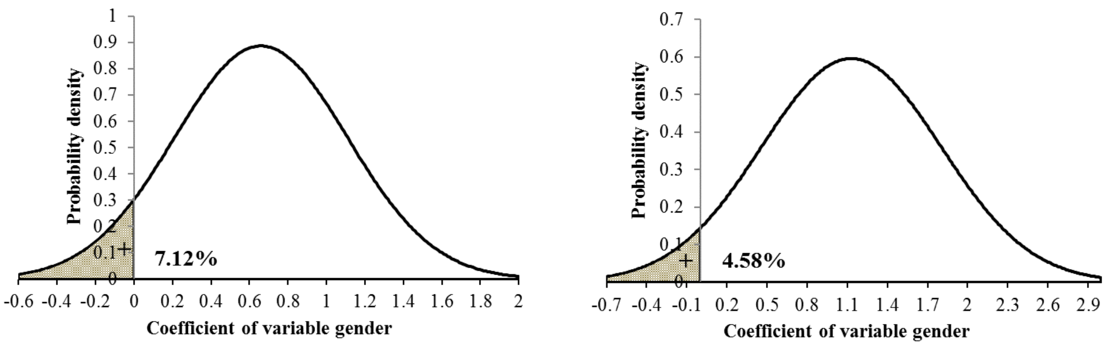
Variables	F/I					NI				
	Mean	Std.	95% Credible Interval		OR	Mean	Std.	95% Credible Interval		OR
			Lower	Upper				Lower	Upper	
Age										
Young vs. Old	0.89	0.077	0.72	1.12	2.47	0.45	0.054	0.31	0.57	1.57
Middle-aged vs. Old	0.36	0.056	0.23	0.49	1.45	0.38	0.043	0.28	0.48	1.44
Gender										
Male vs. Female	0.66	0.024	0.56	0.74	1.88	1.13	0.078	0.92	1.32	2.88
Std. dev. of parameter distribution	0.45	0.033	0.32	0.63		0.67	0.172	0.32	1.11	
E-bike Behavior										
Violation vs. Normal riding	1.21	0.443	0.31	2.12	3.24	1.03	0.271	0.47	1.66	2.78
Distraction vs. Normal riding	0.97	0.321	0.32	1.63	2.56	0.77	0.225	0.31	2.27	2.04
Std. dev. of parameter distribution	0.63	0.232	0.16	1.14		0.63	0.254	0.11	1.21	
License Plate or Not										
Yes vs. No	−0.55	0.220	−0.98	−0.11	0.57	−0.53	0.230	−0.98	−0.02	0.59
Bicycle Type										
E-bike vs. E-scooter	−0.44	0.124	−0.67	−0.18	0.65	−0.47	0.156	−0.81	−0.14	0.61
Std. dev. of parameter distribution	0.25	0.067	0.11	0.39		0.32	0.077	0.14	0.52	
Location										
Intersection vs. Segment	0.72	0.253	0.20	1.12	2.03	0.68	0.255	0.17	1.19	1.87
Speed Limit										
≥45 km/h vs. <45 km/h	0.57	0.088	0.37	0.75	1.79	0.62	0.123	0.35	0.87	1.88
Std. dev. of parameter distribution	0.44	0.121	0.17	0.63		0.53	0.117	0.28	0.77	
DIC	423.6									

As shown in Tables 2 and 3, seven risk factors are found to be significantly associated with e-bike collision severities, including age, gender, e-bike behavior, license plate, bicycle type, location, and speed limit. The FP-MNL coefficients are similar in terms of magnitude and sign to those from RP-MNL. Furthermore, the parameters of gender, e-bike behavior, bicycle type, and speed limit are found to be random parameters in RP-MNL, presenting significant heterogeneous effects on e-bike collision severities. The DIC for FP-MNL is 457.3 while the DIC for RP-MNL is 423.6, showing a difference in

DIC of 33.7. As pointed out by Spiegelhalter et al. [20], models with DIC values' difference within 2–7 show less support to the higher DIC model. The comparison results indicate that the proposed RP-MNL has better model performance than the FP-MNL, which confirms that accommodating the unobserved heterogeneity across observations could improve the model fit.

5.2. Interpretation of Model Estimates

Given that the RP-MNL outperforms the FP-MNL, it was selected for evaluating the risk factors associated with e-bike collision severities. As shown in Table 3, gender is found to be significantly associated with e-bike collision severities. Males are 1.88 times more likely to be involved in F/I e-bike collisions than females. Moreover, males are found to be 2.88 times more likely to be involved in NI e-bike collisions than females. This finding is consistent with previous studies [10,12] which showed that males are more crash prone while women are more risk averse. Wang et al. [9] also pointed out that male e-bike riders are more likely to be at fault in collisions than female riders. The parameters for this variable is normally distributed with (0.66, 0.45) for F/I collision and (1.13, 0.67) for NI collision. As shown in Figure 1a,b, the parameters distributions indicate that 92.88% of the male e-bike riders prefer to be involved in F/I e-bike collisions than female riders, whereas the remaining 7.12% of male e-bike riders have lower probability of being involved in F/I e-bike collisions than female riders. Similarly, 95.42% of the male riders have higher likelihood in NI e-bike collisions than female riders, while 4.58% of male riders are less likely to be involved in NI e-bike collisions than female riders. The result confirms the heterogeneous effects across individuals.



(a) (b)

Figure 1. Varying effects of gender on e-bike collision severities. (**a**) Distribution of parameter estimates for F/I collision; (**b**) Distribution of parameter estimates for NI collision.

Age is found to be significantly related to e-bike collision severities. According to the OR analysis, young and middle-aged e-bike riders are found to be 2.47 times and 1.45 times more likely to be involved in F/I e-bike collisions than old e-bike riders. Moreover, young and middle-aged e-bike riders are 1.57 times and 1.44 times more likely to be involved in NI e-bike collisions than old e-bike riders. The finding is consistent with those of previous studies [21,22]. Bernhoft and Carstensen [21] found that older e-bike riders are more cautious than younger riders, leading to a lower collision risk. Furthermore, Wu and Liu [22] found that young and middle-aged e-bike riders have higher probability of violating traffic rules than older e-bike riders. However, this finding is inconsistent with Hu et al. [6] and Wang et al. [9]. Hu et al. [6] found that older e-bike riders have greater injury severity than the younger group. Wang et al. [9]. found that older e-bike riders are more likely to be at fault in a collision.

E-bike collision severity is found to be significantly affected by e-bike riders' behaviors. Violation e-bike riders are 3.24 times more likely to be involved in F/I collisions and 2.78 times more likely to be involved in NI collisions. Distracted e-bike riders were 2.56 times more likely to be involved in F/I collisions and 2.04 times more likely to be involved in NI collisions. The result is straightforward because violation and distraction could increase e-bike riders' crash risk. The parameters of this variable (distraction) follow normal distributions with (0.97, 0.63) for F/I collisions and with (0.77, 0.63) for NI collisions. As shown in Figure 2a,b, the distraction could increase the probability of F/I e-bike collisions for 93.82% e-bike riders, while for the other 6.18% e-bike riders, the probability of F/I e-bike collisions decrease. Similarly, distraction increases probability of NI e-bike collisions for 88.92% e-bike riders, while it decreases probability of NI e-bike collisions for 11.08% e-bike riders.

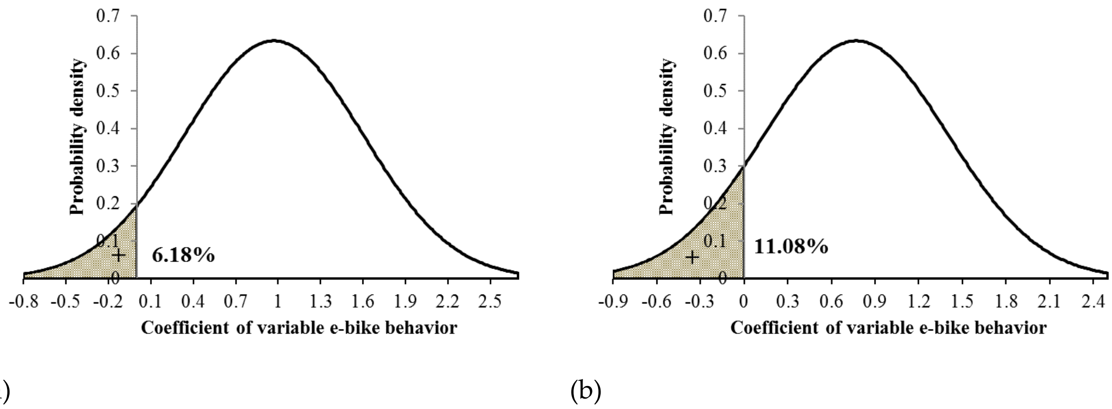
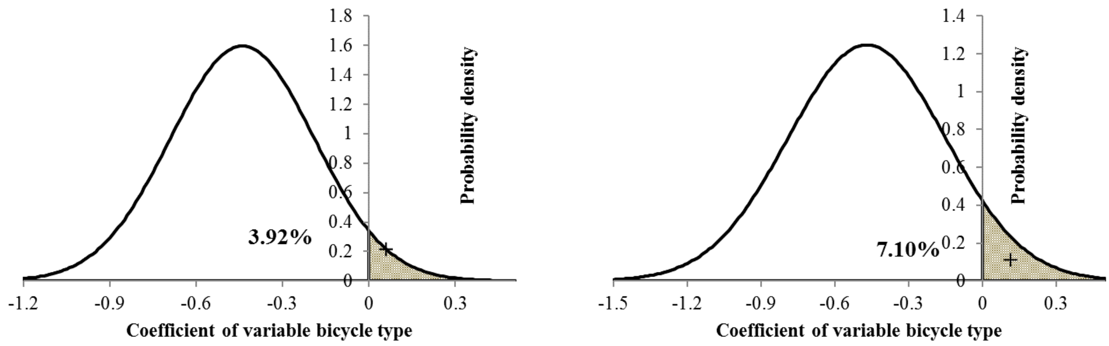


Figure 2. Varying effects of e-bike behavior on e-bike collision severities. (**a**) Distribution of parameter estimates for F/I collision. (**b**) Distribution of parameter estimates for NI collision.

The binary variable license plate is found to be significantly associated with e-bike collision severities. The result showed an OR of 0.57 in F/I e-bike collision and 0.59 in NI e-bike collision for this variable, indicating that e-bike riders who install a license plate on their e-bikes are less likely to be involved in both F/I and NI collisions. The result is consistent with Guo et al. [10] that e-bike collisions have strong negative relations with e-bike license plate use. Weinert et al. [12] highlighted that imposing a license system for e-bikes could make it easier to enforce traffic laws. As such, the encouragement of the use of license plates is an effective countermeasure to improve the e-bike safety.

Bicycle type is significantly related with e-bike collision severities. Based on the OR analysis, e-bikes are found to be 0.65 times less likely to be involved in F/I collisions and 0.61 times less likely to be involved in NI collisions than e-scooters. This finding is in line with several studies [4,16,22] which found that the e-scooters are more risky due to their higher operating speeds and their conflicts with other road users. The parameter of this variable is normally distributed with (−0.44, 0.25) for F/I e-bike collision and (−0.47, 0.32) for NI e-bike collision. As shown in Figure 3a,b, 96.08% of the e-bikes have lower probability of being involved in F/I collisions than e-scooters, whereas the remaining 3.92% of e-bikes have higher probability of being involved in F/I collisions than e-scooters. Moreover, 92.9% of the e-bikes are less likely to be involved in NI collisions than e-scooters, whereas 7.1% of the e-bikes are more likely to be involved in NI collisions than e-scooters. The result implies the heterogeneous effects across bicycle types.

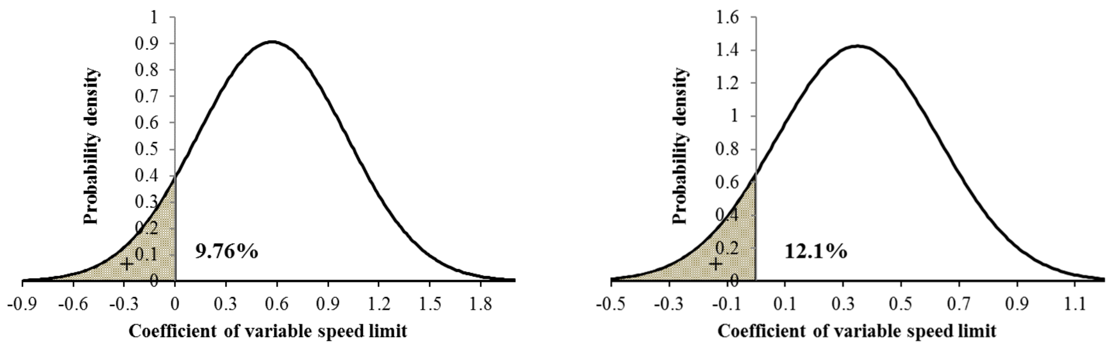


(a) (b)

Figure 3. Varying effects of bicycle type on e-bike collision severities. (**a**) Distribution of parameter estimates for F/I collision. (**b**) Distribution of parameter estimates for NI collision.

Location is found to be a significant risk factor with e-bike collision severities. The OR analysis result shows that F/I e-bike collisions are 2.03 times more likely to occur at intersections than at a road segment. NI e-bike collision are found to be 1.87 times more likely to take place at intersections than at road segment. This finding is intuitive that e-bikes need to handle more complicated driving behaviors, such as interacting with turning vehicles, at intersections than at segments. As such, e-bikes are explored in more traffic conflicts at intersections than at a road segment, which increases the e-bike crash risk.

Speed limit showed significant positive relation with e-bike collision severities. According to the OR result, e-bikes are 1.79 times more likely to be involved in F/I collisions at a speed limit greater than 45 km/h than that at a speed limit less than 45 km/h. As for the NI collisions, e-bikes are found to be 1.88 times more likely to be involved in NI collisions at a speed limit greater than 45 km/h than that at a speed limit less than 45 km/h. The parameters of this variable were found to be random with a normal distribution with (0.57, 0.44) for F/I collision and (0.35, 0.28) for NI collision. The heterogeneous effect of this variable on F/I collisions and NI collisions can be found in Figure 4a,b.



(a) (b)

Figure 4. Varying effects of speed limit on e-bike collision severities. (**a**) Distribution of parameter estimates for F/I collision. (**b**) Distribution of parameter estimates for NI collision.

6. Conclusions

This study investigated the unobserved heterogeneity in severity of e-bike collisions. E-bike collisions data from the city of Ningbo, China were used for the evaluation. A random parameters multinomial logit regression (RP-MNL) was developed to explore the risk factors associated with e-bike collision severities. The RP-MNL was estimated using the full Bayesian approach and compared with the fixed parameters multinomial logit regression (FP-MNL). The unobserved heterogeneous effects associated with observations were captured successfully by the proposed RP-MNL.

The comparison results showed that RP-MNL outperformed FP-MNL reflecting by decreased DIC. The RP-MNL estimates showed that seven risk factors, including age, gender, e-bike behavior, license plate, bicycle type, location, and speed limit, were found to be significantly related to both F/I e-bike collisions and NI e-bike collisions. Although previous studies found risk factors such as age [6,21,22] and gender [9,10,12] contributed to e-bike collisions, they are limited to their capacity to explain the impact of such risk factors on different severity levels of e-bike collisions. In addition, these prior studies did not take the unobserved heterogeneity into the modeling process. In this study, four risk factors, i.e., gender, e-bike behavior, bicycle type, and speed limit, were found to have heterogeneous effects on e-bike collision severities, appearing in the form of random parameters in the statistical model.

There are several limitations in this study. The data used in this study was limited to only one city. E-bike riders' driving patterns may vary among different cultural contexts and traffic environments, leading to different characteristics of e-bike collisions. As such, the transferability of the proposed model should be validated using the data from other cities. Given the difficulty in collecting field data, e-bike riders' socioeconomic characteristics, such as income, educational background, and profession, were not included in the model. Future work could include these variables in the model and explore the mechanism of these variables on e-bike collision severities. Due to the difference in fatal injuries among age groups, future study could model e-bike crash severity separately for different age groups.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design, Y.G., Y.W., J.L., J.Z.; Data collection, J.Z.; Analysis and interpretation of results, Y.G., Y.W., J.L., J.Z.; Draft manuscript preparation, Y.G., Y.W., J.L., J.Z. All authors reviewed the results and approved the final version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (71701046), the China Postdoctoral Science Foundation funded project (2017M571644; 2018T110427), the Fundamental Research Funds for the Central Universities (YBJJ1533), the Scientific Innovation Research of College Graduates in Jiangsu Province (KYLX_0173), Natural Science Foundation of Zhejiang Province (No. LQ19E080003), and Natural Science Foundation of Ningbo Municipality (No. 2018A610127).

Conflicts of Interest: The authors have declared that no competing interests exist.

References

1. Guo, Y.; Li, Z.; Wu, Y.; Xu, C. Exploring unobserved heterogeneity in bicyclists' red-light running behaviors at different crossing facilities. *Accid. Anal. Prev.* **2018**, *115*, 118–127. [CrossRef] [PubMed]
2. Parker, A.A. Electric power-assisted bicycles reduce oil dependence and enhance the mobility of the elderly. In Proceedings of the 29th Australasian Transport Research Forum (ATRF), Melbourne, Australia, 27–29 September 2006.
3. Feng, Z.; Raghuwanshi, R.P.; Xu, Z.; Huang, D.; Zhang, C.; Jin, T. Electric-bicycle-related injury: A rising traffic injury burden in China. *Inj. Prev.* **2010**, *16*, 417–419. [CrossRef] [PubMed]
4. Yao, L.; Wu, C. Traffic safety for electric bike riders in China: Attitudes, risk perception, and aberrant riding behaviors. *Transp. Res. Rec. J. Transp. Res. Board* **2012**, *2314*, 49–56. [CrossRef]
5. Lawinger, T.; Bastian, T. Neue Formen der Zweiradmobilität. Eine empirische Tiefenanalyse von Pedelec-Unfällen in Baden-Württemberg. *Zeitschrift für Verkehrssicherheit* **2013**, *59*, 99–106.
6. Hu, F.; Lv, D.; Zhu, J.; Fang, J. Related risk factors for injury severity of e-bike and bicycle crashes in Hefei. *Traffic Inj. Prev.* **2014**, *15*, 319–323. [CrossRef] [PubMed]
7. Weber, T.; Scaramuzza, G.; Schmitt, K.-U. Evaluation of e-bike accidents in Switzerland. *Accid. Anal. Prev.* **2014**, *73*, 47–52. [CrossRef] [PubMed]
8. Du, W.; Yang, J.; Powis, B.; Zheng, X.; Ozanne-Smith, J.; Bilston, L.; Wu, M. Understanding on-road practices of electric bike riders: An observational study in a developed city of China. *Accid. Anal. Prev.* **2013**, *59*, 319–326. [CrossRef] [PubMed]

9. Wang, C.; Xu, C.; Xia, J.; Qian, Z. Modeling faults among e-bike-related fatal crashes in China. *Traffic Inj. Prev.* **2017**, *18*, 175–181. [CrossRef] [PubMed]
10. Guo, Y.; Zhou, J.; Wu, Y.; Chen, J. Evaluation of factors affecting e-bike involved crash and e-bike license plate use in China using a bivariate probit model. *J. Adv. Transp.* **2017**. [CrossRef]
11. Mannering, F.L.; Shankar, V.; Bhat, C.R. Unobserved heterogeneity and the statistical analysis of highway accident data. *Anal. Methods Accid. Res.* **2016**, *11*, 1–16. [CrossRef]
12. Weinert, J.X.; Ma, C.; Yang, X.; Cherry, C.R. Electric two-wheelers in China: Effect on travel behavior, mode shift, and user safety perceptions in a medium-sized city. *Transp. Res. Rec.* **2007**, *2038*, 62–68. [CrossRef]
13. Otte, D.; Facius, T.; Mueller, C. Pedelecs im Unfallgeschehen und Vergleich zu konventionellen nicht motorisierten Zweiraedern. *VKU Verkehrsunfall und Fahrzeugtechnik* **2014**, *2*, 48–60.
14. Papoutsi, S.; Martinolli, L.; Braun, C.T.; Exadaktylos, A.K. E-bike injuries: Experience from an urban emergency department-A retrospective study from Switzerland. *Emerg. Med. Int.* **2014**, *2014*, 850236. [CrossRef] [PubMed]
15. Schepers, J.P.; Fishman, E.; Den Hertog, P.; Wolt, K.K.; Schwab, A.L. The safety of electrically assisted bicycles compared to classic bicycles. *Accid. Anal. Prev.* **2014**, *73*, 174–180. [CrossRef] [PubMed]
16. Guo, Y.; Liu, P.; Bai, L.; Xu, C.; Chen, J. Red light running behavior of electric bicycles at signalized intersections in China. *Transp. Res. Rec. J. Transp. Res. Board* **2014**, *2468*, 28–37. [CrossRef]
17. Guo, Y.; Li, Z.; Wu, Y.; Xu, C. Evaluating factors affecting electric bike users' registration of license plate in China using Bayesian approach. *Transp. Res. Part F Traffic Psychol. Behav.* **2018**, *59*, 212–221. [CrossRef]
18. Guo, Y.; Osama, A.; Sayed, T. A cross-comparison of different techniques for modeling macro-level cyclist crashes. *Accid. Anal. Prev.* **2018**, *113*, 38–46. [CrossRef] [PubMed]
19. Washington, S.P.; Karlaftis, M.G.; Mannering, F. *Statistical and Econometric Methods for Transportation Data Analysis*, 2nd ed.; Taylor & Francis Group: Boca Raton, FL, USA, 2010; pp. 200–230. ISBN 978-1-4200828-5-2.
20. Spiegelhalter, D.J.; Best, N.G.; Carlin, B.P.; Van Der Linde, A. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B* **2002**, *64*, 583–639. [CrossRef]
21. Bernhoft, I.M.; Carstensen, G. Preferences and behaviour of pedestrians and cyclists by age and gender. *Transp. Res. Part F Traffic Psychol. Behav.* **2008**, *11*, 83–95. [CrossRef]
22. Wu, O.; Liu, Q.M. Electric bicycle related injury and risk factors in Hangzhou. *J. Environ. Occup. Med.* **2012**, *9*, 1–17. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).