Examining the injury severity of public bus–taxi crashes: A random parameters logistic model with heterogeneity in means approach

Qiang Zeng[a,b], Zikang Li[a], Qianfang Wong[a], S.C. Wong[c], Pengpeng Xu[a,d*]

[a] School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, Guangdong, 510641, China

[b] Key Laboratory of Highway Engineering of Ministry of Education, Changsha University of Science & Technology, Changsha, Hunan, 410114, China

[c] Department of Civil Engineering, The University of Hong Kong, Hong Kong, 999077, China

[d] Hunan Key Laboratory of Smart Roadway and Cooperative Vehicle-Infrastructure Systems, Changsha University of Science & Technology, Changsha, Hunan, 410114, China

* *Corresponding author.* E-mail address: pengpengxu@yeah.net

Funding: This research is jointly supported by the National Natural Science Foundation of China (No. 52302433), Guangdong S&T Programme (No. 2024A1111120009), a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (No. 17202824), Natural Science Foundation of Guangdong Province, China (No. 2023A1515012404; No. 2024A1515011578), Science and Technology Projects in Guangzhou (No. 2024A04J3838), Open Fund of the Key Laboratory of Highway Engineering of Ministry of Education, Changsha University of Science & Technology (No. kfj230301), and Open Fund of Hunan Key Laboratory of Smart Roadway and Cooperative Vehicle-Infrastructure Systems, Changsha University of Science & Technology (No. kfj230801). Prof. S.C. Wong was supported by the Francis S Y Bong Professorship in Engineering.

Examining the injury severity of public bus–taxi crashes: A random parameters logistic model with heterogeneity in means approach

ABSTRACT

Public buses and taxis play crucial roles in urban transportation. Ensuring their safety is of paramount importance to develop sustainable communities. This study investigated the significant factors contributing to the injury severity of bus–taxi crashes, using the crash data recorded by the police in Hong Kong from 2009 to 2019. To account for the unobserved heterogeneity, the random parameters logistic model with heterogeneity in means was elaborately developed. The results revealed that taxi driver age, bus age, traffic congestion, and taxi driver behavior had significantly heterogeneous effects on the injury severity of bus-taxi crashes and that the mean value of the random parameter for severe traffic congestion was likely to increase if the taxi's age was less than five years. Taxi driver gender, rainfall, time of day, crash location, bus driver behavior, and collision type were found to significantly affect the bus–taxi crash severity. Specifically, female taxi drivers, old taxis, rainfall, midnight, improper manipulation of bus and taxi drivers, head-on and sideswipe collision types, and non-intersections were associated with a higher likelihood of fatal and severe crashes. Based on our findings, targeted countermeasures were proposed to mitigate the injury severity of bus–taxi crashes.

Keywords: Bus–taxi crashes; Injury severity; Unobserved heterogeneity; Random parameters; Heterogeneity in means

1. Introduction

Public transit bus and taxi are two main types of commercial motor vehicles which provide daily travel services for a mass of urban residents, especially for those in Asian megacities. For example, in Hong Kong almost 40% of daily trips are made by public transit buses or taxis (Meng et al., 2017; Zhou et al., 2020). Thus, their safety issues have always been a primary concern for the public and transportation agencies. Probably due to their high exposure on roadways, the crash rates (measure by the average crash per vehicle) of buses and taxis are usually higher than that of private cars (Ma et al., 2019; Peng et al., 2022, 2024), despite that bus is deemed as the safest transport mode in many developed countries and regions. In addition, buses are often loaded by more passengers than other motor vehicles. Once a bus was involved in a traffic accident, onboard passengers might suffer from injury, thereby resulting in severe adverse outcomes (Shen et al., 2022). Besides, according to the statistics from the Hong Kong Police Force, nearly 20% of the taxi crashes involved buses, while approximately 58% of the bus crashes involved taxis during the last decade. Such a crash type is thereby worth of particular investigation.

In the extant literature, a number of studies have investigated the relationship between the injury severity of bus crashes and driver characteristics (e.g., age, gender, use of safety equipment, and driving behavior), vehicle features (e.g., type, age, and seating capacity), roadway design (e.g., number and width of lanes, shoulder width, and speed limit), traffic conditions, environment (e.g., peak or off-peak hours, weather, and lighting conditions), and crash configuration (e.g., collision type and number of vehicles involved). Meanwhile, there is a limited number of studies (Meng et al., 2017) which have examined the potential factors contributing to the injury severity of taxi crashes. Nonetheless, to the best of our knowledge, there is no published work which have analyzed the injury severity of crashes between bus and taxi. Exploring the effects of driver and vehicle factors related to bus and taxi on the crash injury severity is beneficial to design effective countermeasures for preventing passenger injuries on public transit, which is very important for promoting its attractiveness and thereby building a sustainable community.

To bridge the research gap, this research focuses on examining the injury severity of two-vehicle crashes between urban transit bus and taxi. A bus–taxi crash dataset from Hong Kong is used for the empirical analysis, and a random parameters logistic model with heterogeneity in means is developed to link the crash severity and the observed factors.

The remainder of this study is organized as follows. Section 2 reviews the existing literature on the analysis of bus and taxi crash severity. The collected crash data are described in Section 3. The structure of the crash severity model is introduced in Section 4. The model results are presented and analyzed in Section 5. Section 6 discusses the practical implications of the major findings. Section 7 concludes the research and gives some future research directions.

2. Literature review

2.1. Bus crash severity analysis

Many developing countries are faced with serious problems regarding bus transit safety. Several safety researchers from these countries have analyzed the bus crash severity. For example, Barua and Tay (2010) investigated the injury severity of transit bus crashes in Dhaka, Bangladesh, using a seven-year crash data. They found that fatal bus crashes were more likely to occur on locations without police control, during off-peak periods, on weekends, and with the involvement of vulnerable road users. Based on the analysis of transit bus crash severity in Mashhad, Iran, Nasri and Aghabayk (2020) concluded that a high speed limit and the involvement of vulnerable road users would significantly increase the injury severity of bus crashes. Focusing on the severity of bus/minibus crashes in Ghana, Sam et al. (2018) found that drunk driver, the absence of traffic control or roadway median, adverse roadway condition, the involvement of pedestrians, nighttime, and weekend were associated with severe bus crashes. Based on the research conducted by Sam et al. (2018), Tamakloe et al. (2021) further stratified the bus/minibus crashes into different groups via day of week (weekday or weekend), roadway pavement (good or bad), and lighting conditions (daylight, dark-lighted, or dark), and examined the factors contributing to the severity of bus/minibus crashes in each group. They found that the effects of certain factors (such as those related to vehicle maneuver) were significantly different across the bus crash groups.

While the transit bus is regarded as a safe mode for passenger transport in developed countries and regions, the bus safety issues have also raised the public concerns, since the annual casualties caused by bus crashes are not negligible. Chimba et al. (2010) investigated the injury severity of 4,528 bus crashes in Florida, US, and found that higher traffic volumes, wider lanes, and wider shoulders would increase the injury severity. Based on the five-year crash data from the General Estimates System in the US, Kaplan and Prato (2012) found that the female, young, and old drivers, inattentive and risky driving behavior, very low and very high speed limits were more likely to result in severe bus crash injuries. In addition, Feng et al. (2016) investigated the factors contributing to fatal bus crashes in the US. Prato and Kaplan (2014) examined the bus crash severity and passenger injury in Denmark, and found that red-light violation was significantly linked to a higher bus crash severity and an increased probability of passenger injury. Shen et al. (2022) developed an extended hierarchical ordered probit model with heteroscedasticity for analyzing the bus crash severity in the UK. The results indicated that significant heteroscedasticity existed in the effects of certain factors, such as roadway surface condition and speed limit. Yoon et al. (2017) examined the effects of regional characteristics on the injury severity of local bus crashes in South Korea, and found that the crash severity was significantly associated with some regional characteristics (including emergency medical environment and ratio of elderly in the community). Besides, Zhou et al. (2020) conducted a comparative analysis of bus passenger injury severity in collisions and non-collisions in Hong Kong. Samerei et al. (2021a, 2021b) introduced the association rule discovery method to identify the factors affecting the injury severity of bus/bus-pedestrian crashes in Victoria, Australia.

Most of the aforementioned studies have investigated the crash severity of urban transit buses, while some other studies focus on the crash severity of intercity buses. For example, Chu (2014) examined the injury severity of high-deck bus crashes on freeway, and found that human factors, including driver fatigue, reckless/drunk driving, and not wearing a seat belt, were significantly associated with severe crash outcomes. Rahman et al. (2011) analyzed the injury risk of crashes involving various bus types (including school bus, transit bus, and intercity bus) and categorized the bus crashes into four groups based on the location (highway or not) and the number (one or two) of involved vehicles. They found that weather conditions had significant effects on the injury severity of crashes in all groups.

2.2. Taxi crash severity analysis

As personalized transport service providers, taxis' safety problem has also raised considerable public concerns. In the literature, there is a quantity of works (Chin and Huang, 2009; Peng et al., 2022; Tay and Choi, 2016; Truong et al., 2020; Wang et al., 2014, 2021), which have investigated the effects of taxi drivers' demographic attributes (e.g., age, gender, and income), driving behavior and attitudes, working conditions, and psychological conditions on their involvement of traffic crashes. Ma et al. (2019) conducted a spatial analysis of taxi crash frequency at the census-tract level, and revealed the effects of road networks, land use, and demographic-economic characteristics on taxi crash risk.

However, there are only a few studies on the injury severity of taxi crashes. Specifically, Lam et al. (2004) investigated the environmental factors contributing to the fatality and injury of taxi drivers in traffic crashes, and found that night time and carrying passengers significantly increased their fatality and injury risk. Meng et al. (2017) conducted an occupant-level analysis of taxi crash injury severity. The results indicated that driver age, number of vehicles involved, crash time, and vehicle impact points had significant effects on the injury severity levels for both taxi drivers and passengers, while passenger age and gender had significant effects on the injury severity of taxi passengers only. Chung (2018) analyzed injury severity of taxi-pedestrian crashes based on a reconstructed crash dataset with information derived from vehicle black boxes. The results indicated that the crash severity was significantly associated with some factors during the crash occurrence (such as the first/second/third impact region of pedestrians). More recently, Chung (2022) incorporated the data from the in-vehicle recording devices into the injury severity analysis of crashes between taxis and two-wheelers, and found that the collision speed of taxis and collision types had significant effects on the injury severity. Besides, Park et al. (2017) investigated the effects of company-level and regional-level factors on the crash severities of various commercial motor vehicles (including bus, taxi, and large truck), with controlling the factors related to drivers and vehicles.

3. Data

The data for this study were sourced from the Traffic Road Accident Database System, which is maintained by the Hong Kong Police Force and the Hong Kong Transport Department. This database includes the detailed information on casualties, vehicle involvement, and accident environment, which was recorded by the police officers at the crash scenes (Meng et al., 2017; Xu et al., 2022; Ye et al., 2024; Zeng et al., 2023; Zhai et al., 2019). It is important to note that only accidents resulting in injuries on public roads are included in this database (Xu et al., 2016; 2019). To ensure the reliability and quality of the analysis, we extracted a total of 3,404 transit bus–taxi crashes from the database, by simultaneously retrieving information on the number and type of vehicles involved and excluding samples with incomplete crash information. The process of retrieving and extracting the two-vehicle bus–taxi crashes in Hong Kong from 2009 to 2019 is illustrated in Fig. 1.

The Hong Kong police stratifies the injury severity into three levels: fatal, serious injury, and slight injury. Since the proportion of fatalities in our sample was significantly less than 1%, we combined it with serious injury to constitute a single level, as in many previous studies on Hong Kong traffic safety (Meng et al., 2017; Wang et al., 2019; Xu et al., 2016; Zeng et al., 2023; Zhai et al., 2019; Zhou et al., 2020). Consequently, the dependent variable in the present study became a dichotomous outcome: KSI (killed and serious injury) and SI (slight injury).

For each crash, the attributes on driver (e.g., the driver age, gender, and improper behavior of bus and taxi drivers), vehicle (e.g., the vehicle age, insurance, and defect of the bus and taxi), traffic congestion, roadway (e.g., roadway type and speed limit), environment (e.g., weather and lighting conditions), and crash configuration (e.g., crash time and type) were extracted from the database and used as explanatory variables for the analysis of bus-taxi crash severity. The categorization and descriptive statistics of the crash severity and its explanatory variables are presented in Table 1.

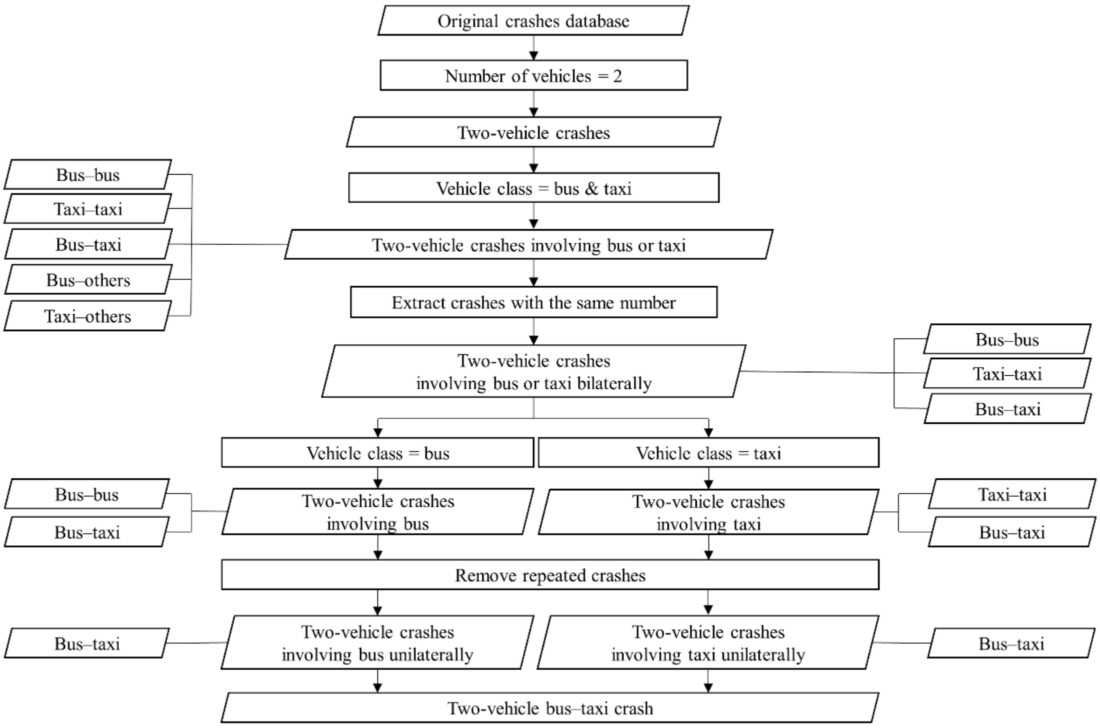

Fig. 1. The flowchart used to extract two-vehicle taxi–bus crashes

Table 1. Descriptive statistics of variables included in the models.

Variables	Number	KSI (%)	SI (%)
Taxi–bus crash severity	3404	305 (8.96%)	3099 (91.04%)
Year of crashes			
2009*	291	26 (8.93%)	265 (91.07%)
2010	305	25 (8.20%)	280 (91.80%)
2011	301	27 (8.97%)	274 (91.03%)
2012	313	26 (8.31%)	287 (91.69%)
2013	321	37 (11.53%)	284 (88.47%)
2014	311	27 (8.68%)	284 (91.32%)
2015	331	31 (9.37%)	300 (90.63%)
2016	320	29 (9.06%)	291 (90.94%)
2017	305	29 (9.51%)	276 (90.49%)
2018	305	23 (7.54%)	282 (92.46%)
2019	301	25 (8.31%)	276 (91.69%)
Time of day			
07:00–10:59	736	60 (8.15%)	676 (91.85%)
11:00–16:59*	919	71 (7.73%)	848 (92.27%)
17:00–20:59	812	71 (8.74%)	741 (91.26%)
21:00–06:59	937	103 (10.99%)	834 (89.01%)
Day of week			
Weekday*	2428	218 (8.98%)	2210 (91.02%)

Weekend	976	87 (8.91%)	889 (91.09%)
Crash location			
Junctions	1278	95 (7.43%)	1183 (92.57%)
Road segments*	2126	210 (9.88%)	1916 (90.12%)
Weather			
Clear*	3175	274 (8.63%)	2901 (91.37%)
Not clear	229	31 (13.54%)	198 (86.46%)
Rainfall			
Rain	479	61 (12.73%)	418 (87.27%)
Not rain*	2925	244 (8.34%)	2681 (91.66%)
Street light			
Adequate*	1744	160 (9.17%)	1584 (90.83%)
Inadequate	1660	145 (8.73%)	1515 (91.27%)
Speed limit			
Less than 50km/h	12	3 (25.00%)	9 (75.00%)
50km/h*	3200	278 (8.69%)	2922 (91.31%)
More than 50km/h	192	24 (12.5%)	168 (87.50%)
Traffic congestion			
Severe	526	38 (7.22%)	488 (92.78%)
Moderate	1101	94 (8.54%)	1007 (91.46%)
None*	1777	173 (9.74%)	1604 (90.26%)
Road type			
One way*	1372	112 (8.16%)	1260 (91.84%)
Two-way	1121	109 (9.72%)	1012 (90.28%)
Dual carriageway	690	63 (9.13%)	627 (90.87%)
More than two carriageways	221	21 (9.50%)	200 (90.50%)
Collision type			
Front	111	25 (22.52%)	86 (77.48%)
Side	300	48 (16.00%)	252 (84.00%)
Rear-end*	1185	100 (8.44%)	1085 (91.56%)
Scrape	1218	81 (6.65%)	1137 (93.35%)
Other collision	590	51 (8.64%)	539 (91.36%)
Taxi driver age			
< 45	496	52 (10.48%)	444 (89.52%)
45–59	1780	143 (8.03%)	1637 (91.97%)
≥ 60*	1128	110 (9.75%)	1018 (90.25%)
Taxi driver gender			
Male*	3349	298 (8.90%)	3051 (91.10%)
Female	55	7 (12.73%)	48 (87.27%)
Taxi age			
< 5	833	57 (6.84%)	776 (93.16%)
5–9	907	82 (9.04%)	825 (90.96%)
≥ 10*	1664	166 (9.98%)	1498 (90.02%)
Taxi insurance			
Valid*	3226	293 (9.08%)	2933 (90.92%)
Invalid	178	12 (6.74%)	166 (93.26%)
Taxi defect			
Defect	117	14 (11.97%)	103 (88.03%)
None*	3287	291 (8.85%)	2996 (91.15%)

Taxi driver factor			
Limited driving space	198	8 (4.04%)	190 (95.96%)
Violation	1309	117 (8.94%)	1192 (91.06%)
Negligence	185	29 (15.68%)	156 (84.32%)
Improperness	30	3 (10.00%)	27 (90.00%)
Others*	1682	148 (8.80%)	1534 (91.2%)
Bus driver age			
< 45	1050	93 (8.86%)	957 (91.14%)
45–59	1678	154 (9.18%)	1524 (90.82%)
≥ 60*	676	58 (8.58%)	618 (91.42%)
Bus driver gender			
Male*	3312	295 (8.91%)	3017 (91.09%)
Female	92	10 (10.87%)	82 (89.13%)
Bus age			
< 5	862	82 (9.51%)	780 (90.49%)
5–9	928	79 (8.51%)	849 (91.49%)
≥ 10*	1614	144 (8.92%)	1470 (91.08%)
Bus insurance			
Valid*	3350	301 (8.99%)	3049 (91.01%)
Invalid	54	4 (7.41%)	50 (92.59%)
Bus defect			
Defect	95	11 (11.58%)	84 (88.42%)
None*	3309	294 (8.88%)	3015 (91.12%)
Bus driver factor			
Limited driving space	406	27 (6.65%)	379 (93.35%)
Violation	880	63 (7.16%)	817 (92.84%)
Negligence	100	20 (20.00%)	80 (80%)
Improperness	55	8 (14.55%)	47 (85.45%)
Others*	1963	187 (9.53%)	1776 (90.47%)

*: reference category; KSI: killed and serious injury; SI: slight injury.

4. Methods

4.1 Model specification

The logistic model serves as the basic statistical method for analyzing crash severity with binary outcomes. Let Y_i denote as the observed injury severity of the bus–taxi crash i. $Y_i = 1$ means that the crash severity is KSI, while $Y_i = 0$ indicates that the crash severity is slight injury. The logistic regression model is formulated as follows (Hosseinzadeh et al., 2021; Xu et al., 2016; Zeng et al., 2023; Zhou et al., 2020):

$$Y_i \sim Binomial(p_i) \tag{1}$$

$$logit(p_i) = ln\left(\frac{p_i}{1-p_i}\right) = \boldsymbol{\beta}\mathbf{x}_i \tag{2}$$

where p_i is the probability of bus-taxi crash i resulting in KSI, $\mathbf{x}_i$ is a vector of explanatory variables observed in crash i, and $\boldsymbol{\beta}$ is a vector of estimable coefficients (including a constant term).

Unobserved heterogeneity is an important characteristic of crash severity data, since there may be some unobserved factors which have significant effects on crash severity and are correlated with the observed factors (Mannering et al., 2016). Several methods, such as the random parameters (Chen et al., 2017, 2019; Saeed et al., 2019; Wang et al., 2024; Xiao et al., 2024), latent classes (Ma et al., 2024; Sun et al., 2024), and Markov switching approaches, have been proposed to account for the unobserved heterogeneity. Among them, the random parameters models are the most widely used (Riccardi et al., 2023; Sarker e al., 2023; Seraneeprakarn et al., 2017; Se et al., 2022; Truong et al., 2020; Wang et al., 2020; Waseem et al., 2019; Xiao et al., 2024; Yu et al., 2020; Yuan et al., 2022; Zhou et al., 2020). Conventional random parameters logistic model is developed by allowing the coefficients to vary across the observations, that is, the $\boldsymbol{\beta}$ in Eq. (2) is transformed to be $\boldsymbol{\beta}_i$, which is formulated as follows:

$$\boldsymbol{\beta}_i = \boldsymbol{b} + \boldsymbol{\varepsilon}_i \tag{3}$$

where $\boldsymbol{\beta}_i$ is a vector of random parameters for the crash i, and $\boldsymbol{b}$ is a vector of fixed coefficients corresponding to the mean of the random parameters. $\boldsymbol{\varepsilon}_i$ is a vector of the random terms, which are usually assumed to be independent and follow normal distributions with zero means for the crash i.

To further account for the potential heterogeneity in the means of random parameters, a random parameters logistic model with heterogeneity in means is developed by formulating the random parameters vector $\boldsymbol{\beta}_i$ as follows (Behnood et al., 2017a; Das et al., 2024; Tamakloe et al., 2023):

$$\boldsymbol{\beta}_i = \boldsymbol{b} + \mathbf{M}\boldsymbol{\omega}_i + \boldsymbol{\varepsilon}_i \tag{4}$$

where $\boldsymbol{\omega}_i$ is a vector of explanatory variables from crash i that affect the means of $\boldsymbol{\beta}_i$, and $\mathbf{M}$ is a vector of estimable parameters corresponding to $\boldsymbol{\omega}_i$.

With crash-specific unobserved heterogeneity allowed, the $\boldsymbol{\beta}_i$ vector has a continuous density function $Prob(\boldsymbol{\beta}_i = \boldsymbol{\beta}) = f(\boldsymbol{\beta}|\boldsymbol{\varphi})$, where $\boldsymbol{\varphi}$ is a vector of parameters characterizing this function (such as the location and scale). In the present study, the normal distribution was considered as the density function. Thus, the probability of KSI for crash i is:

$$p_i = \int \frac{e^{\boldsymbol{\beta}_i \mathbf{X}_i}}{1+e^{\boldsymbol{\beta}_i \mathbf{X}_i}} f(\boldsymbol{\beta}|\boldsymbol{\varphi}) d\boldsymbol{\beta} \tag{5}$$

To evaluate the impact of variables on outcome probability, a conventional approach is to calculate elasticities. However, in the context of injury severity analysis, variables are typically represented as indicator variables, rendering the calculation of regular elasticity due to the non-differentiability of the probability function. A feasible solution is to compute the change in probabilities for each injury severity level when switching the indicator variable from 0 to 1. This measure, known as the direct pseudo-elasticity (Washington et al., 2020), represents the percentage change in probabilities for each injury severity level resulting from switching the indicator variable. For this research, the pseudo-elasticity is calculated as follows:

$$E_{x_{in}}^{p_i} = \frac{p_i(x_{in}=1) - p_i(x_{in}=0)}{p_i(x_{in}=0)} \tag{6}$$

where $E_{x_{in}}^{p_i}$ is the pseudo-elasticity of the nth independent variable on KSI. x_{in} is the value of the nth independent variable in crash i. Since the pseudo-elasticity is calculated for each observation, the average value of the pseudo-elasticity of the whole dataset is usually taken to quantify the influence of the independent variable on the dependent variable.

The random parameters logistic models were estimated using a simulation-based maximum likelihood approach with 600 Halton draws (Washington et al., 2020). The model estimation process was implemented in the econometric software NLOGIT 6.0.

4.2 Model comparison criteria

To validate the outperformance of the random parameters logistic model with heterogeneity in means, we compared it with the fixed parameters logistic model and the conventional random parameters logistic model, using the Akaike Information Criterion (AIC), the classification accuracies for each injury severity and the entire dataset, and likelihood ratio tests.

As one of the most prevalent criteria for assessing model goodness-of-fit, the AIC measures the magnitude of information lost in a model, which is calculated as follows

(Akaike, 1973):

$$AIC = 2K - 2LL(\boldsymbol{\beta}) \tag{7}$$

where K is the number of estimable parameters in a model, and $LL(\boldsymbol{\beta})$ is its log-likelihood at convergence. Generally, a model with a lower AIC value is preferred.

The classification accuracy is an extensively adopted criterion to evaluate the prediction performance of discrete outcome models (He et al., 2024; Tang et al., 2019; Zeng et al., 2023). Given the two injury severity levels in the research, the results can be divided into four groups, based on the combination of predicted and real crash severity levels: (1) the true positive (TP), i.e., both of the predicted and real severity levels are KSI; (2) the false positive (FP), i.e., the predicted severity level is KSI but the real severity level is slight injury; (3) the true negative (TN), i.e., both of the predicted and real severity levels are slight injury; and (4) the false negative (FN), i.e., the predicted severity level is slight injury but the real severity level is KSI. Accordingly, the classification accuracies for KSI, slight injury, and the entire dataset are defined as following, respectively:

$$CA_{KSI} = \frac{n_{TP}}{n_{TP}+n_{FN}} \tag{8}$$

$$CA_{SL} = \frac{n_{TN}}{n_{TN}+n_{FP}} \tag{9}$$

$$CA_{entire} = \frac{n_{TP}+n_{TN}}{n_{TP}+n_{FN}+n_{TN}+n_{FP}} \tag{10}$$

where n_{TP}, n_{FN}, n_{TN}, and n_{FP} are the numbers of TP, FN, TN, and FP cases, respectively.

Likelihood ratio tests are conducted to justify whether the random parameters logistic model with heterogeneity in means is superior to other candidate models (i.e., the logistic model and random parameters logistic model) in analyzing the crash injury severity. The test statistic is defined as (Washington et al., 2020):

$$\chi^2 = -2\big(LL(\boldsymbol{\beta}_o) - LL(\boldsymbol{\beta}_{RPHM})\big) \tag{11}$$

where $LL(\boldsymbol{\beta}_{RPHM})$ is the log-likelihood at convergence of the random parameters logit model with heterogeneity in means, and $LL(\boldsymbol{\beta}_o)$ is the log-likelihood at convergence of the candidate model. The test statistic is chi-square distributed, with the degree of freedom equal to the difference in the number of estimable parameters.

5. Results

In terms of model specification, we performed a correlation test to avoid the inclusion of highly correlated variables. Fig. 2 illustrates a strong correlation between street light and time of day, suggesting that these two variables should not be included in the models simultaneously. Additionally, rainfall and weather displayed a moderate correlation, with an estimated Spearman's correlation coefficient of –0.51. Other variables showed weak collinearity, as their Spearman's correlation coefficients were less than 0.50. We incorporated all the uncorrelated variables and leveraged the backward method for potential variable selection. A model with a lower AIC value was deemed preferable.

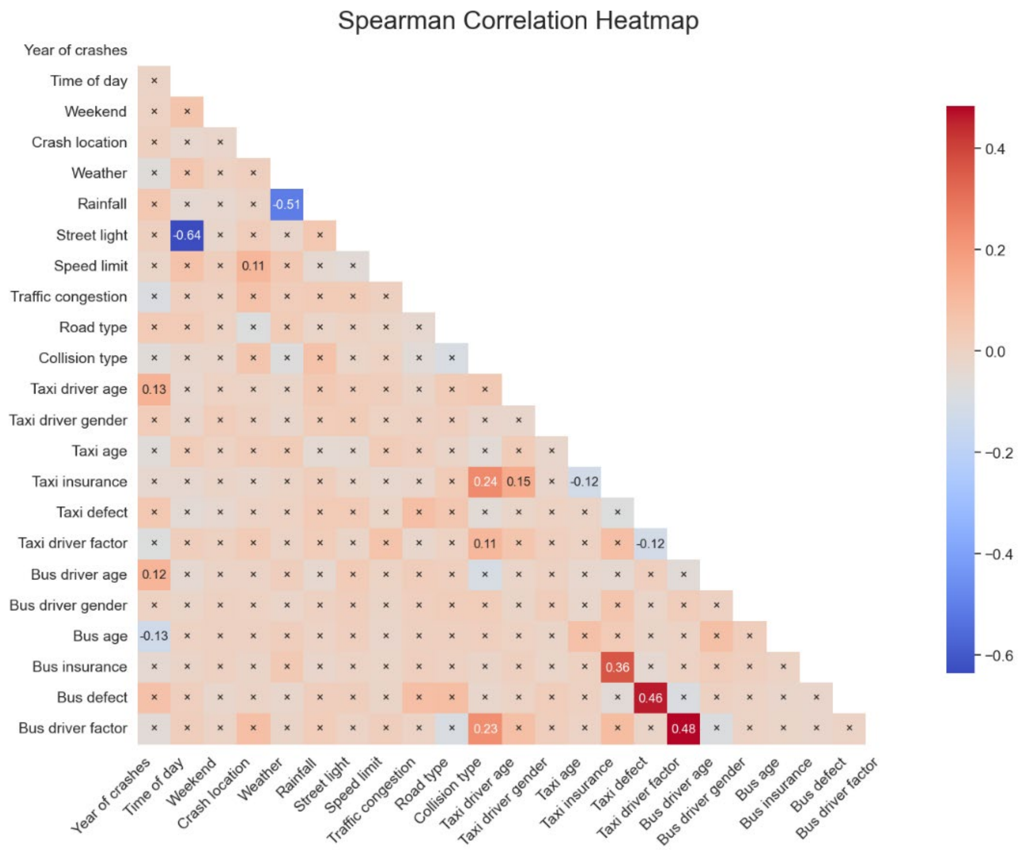

Fig. 2. Results of correlation analysis.

For comparison purpose, in addition to the random parameters logistic model with heterogeneity in means, we developed the fixed parameters logistic model and the standard random parameters model without heterogeneity in means and variances. The performance of these models is presented below, followed by the interpretation of the parameter estimations.

5.1 Model comparison

Table 2 presents the results of parameter estimation and comparison criteria of the logistic model, the random parameters logistic model, and the random parameters logistic model with heterogeneity in means. Only the explanatory variables with significant influence at the 90% confidence level or above were included in the models.

According to the results, the AIC value for the random parameters logistic model with heterogeneity in means was lower than that of the other two models, indicating the better performance of the former model. The classification accuracies of the random parameters logistic model with heterogeneity in means were the highest for each injury severity level and the entire dataset, which further confirms its superiority. Particularly, the classification accuracy for KSI crashes of the random parameters logistic model with heterogeneity in means was about 6.6 times that of the logistic model, and 1.5 times that of the random parameters logistic model. These results confirm previous findings (Alnawmasi and Mannering, 2019; Behnood and Mannering 2017a, 2017b; Se et al., 2022; Waseem et al., 2019) that the specification of a more flexible structure to capture the unobserved heterogeneity can substantially improve the model goodness-of-fit and prediction performance. In addition, the likelihood ratio tests between the random parameters logistic model with heterogeneity in means and the other two models resulted in the p-value less than 0.025, revealing its significant differences from the other alternatives.

5.2 Interpretation of parameter estimations

To explicitly reveal the effects of significant factors on the injury severity of two-vehicle bus–taxi crashes, the estimation results of the random parameters logistic model with heterogeneity in means were interpreted, given its superiority in model overfit.

Table 2. Parameter estimation and comparison criteria results of the bus–taxi crash severity models.

Variables	Logistic model			Random parameters logistic model			Random parameters logistic model with heterogeneity in means			
	Coefficient	SD	*p*	Coefficient	SD	*p*	Coefficient	SD	*p*	Elasticity
Constant	–2.37***	0.43	0.00	–2.24***	0.42	0.00	–2.19***	0.51	0.00	——
Time of day (reference: 11:00–16:59)										
07:00–10:59	–0.03	0.19	0.86	–0.17	0.18	0.34	–0.41*	0.23	0.07	–8.76%
17:00–20:59	0.31	0.21	0.14	–0.16	0.21	0.42	0.53**	0.24	0.03	12.55%
21:00–06:59	0.45**	0.12	0.04	0.60***	0.20	0.00	0.56**	0.25	0.02	15.39%
Crash location: Junctions	–0.40***	0.14	0.01	–0.48***	0.14	0.00	–0.80***	0.18	0.00	–29.80%
Rainfall (Yes=1, No=0)	0.38**	0.18	0.05	0.33*	0.18	0.07	0.64***	0.23	0.00	9.04%
Traffic congestion (reference: none)										
Severe	–0.24	0.20	0.21	–0.53***	0.22	0.01	–0.91***	0.28	0.00	–13.97%
SD (Severe)	——	——	——	1.14***	0.25	0.00	0.76**	0.28	0.02	——
Moderate	–0.10	0.14	0.48	–0.06	0.14	0.64	–0.06	0.17	0.74	–1.85%
Taxi driver age (reference: ≥60)										
< 45	0.12	0.19	0.52	0.21	0.18	0.25	0.11	0.21	0.61	1.54%
45–59	–0.27**	0.12	0.02	–0.14	0.14	0.32	–2.52***	0.28	0.00	–131.26%
SD (45–59)	——	——	——	——	——	——	4.16***	0.30	0.00	——
Taxi driver gender: Female	0.54	0.42	0.20	0.92***	0.35	0.01	1.54***	0.54	0.00	2.48%
Taxi age (reference: ≥10)										
< 5	–0.37**	0.17	0.03	–0.37**	0.17	0.03	–0.77***	0.32	0.01	–18.88%
5–9	–0.10	0.18	0.58	0.02	0.17	0.93	0.02	0.21	0.94	0.44%
Bus age (reference: ≥10)										
< 5	0.14	0.15	0.38	–0.29*	0.15	0.06	–2.36***	0.37	0.00	–59.56%
SD (< 5)	——	——	——	——	——	——	5.67***	0.49	0.00	——
5–9	–0.05	0.15	0.77	–0.06	0.15	0.68	0.06	0.17	0.74	1.59%
Collision type (reference: rear-end)										

Front	0.84***	0.30	0.00	0.71***	0.27	0.01	0.97***	0.40	0.00	3.14%
Side	0.45**	0.22	0.04	0.40	0.24	0.11	0.54*	0.30	0.07	4.71%
Scrape	–0.35***	0.15	0.01	–2.70***	0.35	0.00	–0.41*	0.25	0.10	–14.64%
SD (Scrape)	——	——	——	4.00***	0.35	0.00	——	—	——	——
Others	0.09	0.21	0.66	0.06	0.20	0.76	0.00	0.25	1.00	0.01%
Taxi driver factor (reference: others)										
Limited driving space	–0.93**	0.40	0.02	–0.75**	0.35	0.03	–0.91**	0.40	0.02	–5.29%
Violation	0.33	0.27	0.22	0.24	0.24	0.32	0.74***	0.29	0.01	4.00%
Negligence	–0.05	0.17	0.78	–4.17***	0.48	0.00	–6.15***	0.65	0.00	–235.99%
SD (Negligence)	——	——	——	6.62***	0.56	0.00	9.72***	0.78	0.00	——
Improperness	–0.21	0.65	0.74	–0.26	0.55	0.64	–0.20	0.64	0.75	0.18%
Bus driver factor (reference: others)										
Limited driving space	–0.40	0.27	0.14	–0.25	0.22	0.27	–0.17	0.28	0.54	–2.04%
Violation	0.66**	0.32	0.03	0.52*	0.27	0.06	0.74***	0.29	0.00	2.77%
Negligence	–0.30	0.20	0.13	–0.27	0.18	0.14	–0.37*	0.22	0.10	–9.53%
Improperness	0.67*	0.40	0.09	0.73*	0.39	0.06	0.87**	0.43	0.04	1.69%
Heterogeneity in the mean of the random parameters										
Traffic congestion: Severe (Taxi age < 5 years)		——	——	——	——	——	1.28***	0.41	0.00	——
K			52			57				60
$LL(\beta)$			–977.2			–964.5				–958.3
AIC			2044.0			2043.1				2036.6
CA_{KSI}			12.4%			55.4%				82.6%
CA_{SL}			91.4%			99.9%				99.8%
CA_{entire}			84.2%			95.9%				98.3%
Likelihood ratio tests										
χ^2			23.8			12.4				——
Degree of freedom			8			3				——
p-value			＜0.025			＜0.025				——

Notes: Coeff. = Coefficient; SD = Standard deviation; ***, **, * = Significance level at 1 %, 5 %, 10 %, respectively.

5.2.1 Factors with heterogeneous effects

According to the parameter estimation results of the random parameters logistic model with heterogeneity in means, there were four variables with significantly heterogeneous effects on the bus–taxi crash severity: severe traffic congestion, taxi driver aged 45–59, bus age under five years, and negligent taxi driver, with the random parameters normally distributed. Amongst them, the severe traffic congestion was the only variable with heterogeneity in the mean of its random parameter. Taxi age less than five years was the indicator variable that would increase the mean, making the bus–taxi crashes more likely to result in KSI (relative to the bus–taxi crashes occurring in no traffic congestion). Specifically, as Fig. 3 shows, if the involved taxi's age was equal to or more than five years, the mean of the random parameter was –0.91, with a standard deviation of 0.76. These results indicate that only 11.5% of bus–taxi crashes in severe traffic congestion were more likely to result in KSI. By contrast, when the involved taxi's age was under 5 years, although the standard deviation of the random parameter was invariant, the mean of the random parameter became 0.96, implying that 89.7% of bus–taxi crashes in severe traffic congestion were more likely to result in KSI.

The indicator variable for taxi driver aged 45–59 produced a random parameter with a mean of –2.52 and a standard deviation of 4.16. Compared with crashes involving taxi drivers under the age of 45, 72.77% of the crashes involving drivers aged 45–59 were less likely to result in KSI, as illustrated in Fig. 3. This finding may be attributed to that middle-aged taxi drivers are usually more experienced and conservative during their driving processes (Meng et al., 2017; Peng et al., 2022). Thus, they probably take appropriate actions to mitigate the injuries when colliding to public buses. However, it is important to note that the remaining 27.23% of bus–taxi crashes were more prone to result in KSI incidents, possibly due to the increased vulnerability of older drivers compared to the younger counterparts (Anastasopoulos and Mannering, 2011; Peng et al., 2024; Wu et al., 2016).

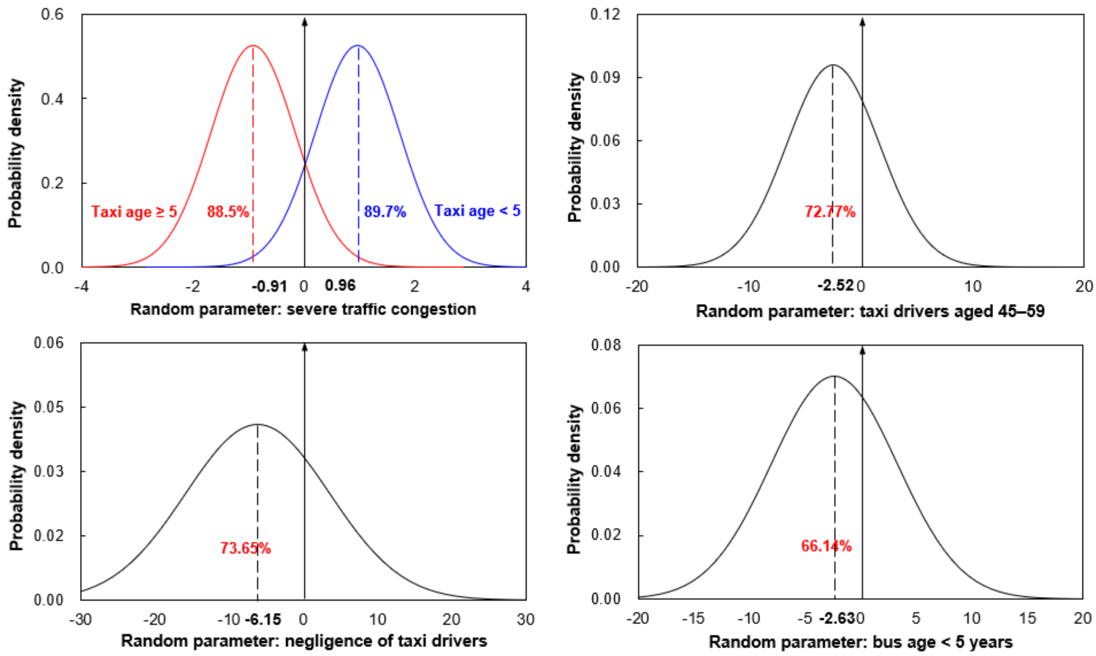

Fig. 3. The distribution of random parameters.

As Fig. 3 shows, the negligence of taxi drivers resulted in a normally distributed parameter with a mean of –6.15 and a standard deviation of 9.72. Given these distributional parameters, 26.35% of the crashes involving taxi driver negligence were more likely to result in KSI. Although this result is somewhat counter-intuitive, one plausible explanation is that as professional drivers, the majority of the distracted taxi drivers subconsciously take self-protective actions to mitigate the potential injuries during the occurrence of bus–taxi crashes.

The random parameter for bus age less than five years yielded a normal distribution with a mean of –2.36 and a standard deviation of 5.67. Given these distributional parameters, the majority (66.14%) of the involved buses with age less than five years were less likely to experience KSI crashes when colliding with taxis. Since relatively new vehicles are deployed with better physical conditions and more advanced safety equipment, this result is reasonable and consistent with the findings of existing studies (Seraneeprakarn et al., 2017; Zeng et al., 2016).

5.2.2 Driver Attributes

The gender of taxi drivers had a significant effect on the crash severity. According to the estimation results, the female taxi driver would increase the KSI probability of a bus–taxi crash by 2.48%, compared with the male taxi driver. The finding is generally consistent with many previous studies (Wu et al., 2016; Zeng et al., 2016). Yu et al. (2020) argued that it may be attributed to the differences between male and female drivers in their driving behavior and physiological response during the process of crash occurrence.

According to the results in Table 2, the taxi driver factors and bus driver factors were important contributors to the injury severity of bus–taxi crashes. Specifically, violation of traffic regulations by taxi and bus drivers was expected to increase the probability of KSI crashes by 4.00% and 2.77%, respectively. These findings generally confirm to empirical justifications and the extant literature (Mallia et al., 2015; Park et al., 2019; Peng et al., 2024) that failure to comply with traffic rules significantly increases the likelihood and severity of crashes. It is also interesting to find that the limited driving space for taxi drivers was negatively related to KSI crashes. One plausible explanation is that drivers tend to slow down in a limited driving space. A lower speed would result in less severe injury sustained by all occupants involved in the collision (Chung, 2022; Zeng et al., 2016).

5.2.3 Vehicle Attributes

In addition to the significant influence on the variance of the normal distribution for severe traffic congestion, taxi age also had a direct effect on crash severity. Specifically, compared with the old taxis (those over nine years), the probability of taxis under five years resulting in KSI crashes would be decreased by 8.88%. This reduction can be attributed to that relatively new vehicles are associated with fewer mechanical failures and are more likely to deploy advanced driver assistance systems (Huang et al., 2011; Zeng et al., 2016). Conversely, older vehicles may exhibit poor responsiveness, particularly when sudden and hard braking maneuvers are conducted (Seraneeprakarn et al., 2017). It is thus not surprising that older taxis were associated with a higher likelihood of KSI crashes.

5.2.4 Roadway Attributes

Crash location was found to have a significant effect on the injury severity of bus–taxi crashes. Specifically, bus–taxi crashes at/near junctions were less likely to result in severe outcomes. The KSI probability of bus–taxi crashes at/near junctions was expected to be lower than those on roadway sections by 29.80%. This finding is expected, because drivers are usually more vigilant with a lower vehicle speed when approaching intersections (Zhou et al., 2020). Additionally, there is often stricter traffic control and law enforcement at junctions, enabling drivers to make more accurate judgments and to take appropriate actions (Sam et al., 2018).

5.2.5 Environment Attributes

Regarding the environment attributes, rainfall was significantly associated with the severity of bus–taxi crashes. On rainy days, the probability of bus–taxi crashes resulting in KSI would increase by 9.04%. The finding is reasonable and in line with many previous studies (Naik et al., 2016; Zhai et al., 2019; Zhou et al., 2020). Precipitation usually makes the roadway surface wet and slippery, increasing the difficulties of vehicle manipulation. The sight distance of drivers may also be reduced by heavy rains. When encountering hazards, longer perception time and braking distance would be required.

5.2.6 Crash configurations

Regarding the time of day, the estimation results indicate that bus–taxi crashes occurring during 17:00–20:59 and 21:00–6:59 were prone to result in KSI. Specifically, compared with 11:00–16:59, the periods 17:00–20:59 and 21:00–6:59 were associated with a 12.55% and 15.39% higher KSI probability, respectively. These findings may be due to the worse visual conditions and decreased alertness of drivers during night (He et al., 2024; Zeng et al., 2019; Zhai et al., 2019). Speeding and fatigue driving are also more likely to occur at midnight, which probably lead to severe outcomes if a crash occurs (Zhou et al., 2020).

As expected, bus–taxi crash severity was significantly associated with collision type. According to the estimated pseudo-elasticities, compared with rear-end collisions, the KSI probability of frontal collisions and side collisions increased by 3.14% and 4.71%, respectively, while the KSI probability of scrape collisions was lower by 14.64%. These findings are also consistent with previous studies (Huang et al., 2011; Zeng et al., 2019) that frontal collisions are the most dangerous, and scrape collisions are usually the safest.

6. Practical implications

Based on our research findings, tailored countermeasures can be formulated to improve the safety of public transportation systems. Specifically, the 4E strategies, namely engineering, enforcement, emergency, and education, are suggested to mitigate the injury severity of crashes between taxis and buses:

- Our analysis results indicate that old and female taxi drivers, along with the illegal and improper driving behaviors of taxi and bus drivers were linked to higher crash severity levels. Thus, the safety awareness of taxi and bus drivers should be enhanced through safety education and publicity activities to ensure the safe driving of these professional drivers, especially the old and female taxi drivers.
- New taxis and buses were less likely to be involved in severe crashes. To improve the safety of old taxis and buses, regular and high-standard maintenances along with the deployment of advanced driving assistant systems are recommended. Besides, if local authorities shorten the legal operation years of buses and taxis, the safety performance of public transit would also be improved.
- To improve the safety performance of public transits at night, traffic facilities such as the street lamps, reflective curb lines, and night guidance signs are suggested to be installed at streets without illumination. Optimizing the schedules of bus and taxi drivers to reduce fatigue driving is also beneficial.
- Reconfigurations of bus and taxi stops to reduce their potential conflicts may alleviate the injury severity of bus–taxi crashes on roadway segments.

- Development and deployment of a proactive traffic management system that integrates both real-time traffic and weather information is recommended to reduce severe bus–taxi crashes in adverse traffic and weather conditions.

7. Conclusions

As two main modes for urban public transportation, the safety of taxis and buses has long been a primary concern of the public. This research identified the factors with significant influence on the injury severity of two-vehicle bus–taxi crashes, based on the police-reported accident data in Hong Kong during a period of 11 years. To address the unobserved heterogeneity, the random parameters logistic model with heterogeneity in means was developed to analyze the effects of the observed factors related to drivers of buses and taxis, vehicles, roadway, traffic, environment, and crash configurations on the injury severity, which was divided into KSI and slight injury.

The results of model comparison indicate that the random parameters logistic model with heterogeneity in means considerably outperformed the fixed parameters logistic model and the standard random parameters logistic model, in term of both model goodness-of-fit and prediction performance. The parameter estimation results show that there were significant heterogeneities in the effects of variables including severe traffic congestion, taxi driver aged 45–59, negligence of taxi drivers, and bus age under five years. In addition, the mean of the random parameter for severe traffic congestion was affected by taxi age. A higher probability of KSI in bus–taxi crashes was associated with female taxi drivers, violation of taxi drivers, violation and improper manipulation of bus drivers, old taxis, rainfall, night and before dawn, front and side collisions, and non-intersection locations. Based on the findings, tailor-made countermeasures were suggested to alleviate the injury severity levels of bus–taxi crashes.

Overall, the present study provides a comprehensive analysis of the injury severity of bus–taxi crashes in Hong Kong. Nonetheless, some enhancements may be required in the future study. For example, inclusion of property-damage-only crashes and examination of the individual injury severities of passengers on buses or taxis are expected to provide deeper insights into their injury mechanisms. In addition, although we have included a host of variables during model specification, several potential risk factors (e.g., collision speed) are unavailable. Future studies towards integration of police reports with other data sources, such as the in-vehicle videos (Chung, 2022; Wang et al., 2024), are highly advocated. Methodologically, the random parameters logistic model with heterogeneity in means was developed, because the heterogeneous effects in the variances of the random parameters were found not to be statistically significant. More sophisticated models, such as the correlated random parameters logistic model (Ahmed et al., 2021), can be leveraged to better accommodate the unobserved heterogeneity.

Disclosure statement: The authors report there are no competing interests to declare.

References

Ahmed, S.S., Cohen, J., Anastasopoulos, P.Ch. (2021). A correlated random parameters with heterogeneity in means approach of deer-vehicle collisions and resulting injury-severities. *Analytic Methods in Accident Research*, 30, 100160.

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, 60(2), 255–265.

Alnawmasi, N., Mannering, F. (2019). A statistical assessment of temporal instability in the factors determining motorcyclist injury severities. *Analytic Methods in Accident Research*, 22, 100090.

Anastasopoulos, P.C., Mannering, F.L. (2011). An empirical assessment of fixed and random parameter logit models using crash-and non-crash-specific injury data. *Accident Analysis & Prevention*, 43(3), 1140–1147.

Barua, U., Tay, R. (2010). Severity of urban transit bus crashes in Bangladesh. *Journal of Advanced Transportation*, 44(1), 34–41.

Behnood, A., Mannering, F. (2017a). The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. *Analytic Methods in Accident Research*, 14, 41–53.

Behnood, A., Mannering, F. (2017b). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research*, 16, 35–47.

Chen, S., Saeed, T.U., Labi, S. (2017). Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. *Analytic Methods in Accident Research*, 16, 75–89.

Chen, S., Saeed, T.U., Alqadhi, S.D., Labi, S. (2019). Safety impacts of pavement surface roughness at two-lane and multi-lane highways: Accounting for heterogeneity and seemingly unrelated correlation across crash severities. *Transportmetrica A: Transport Science*, 15(1), 18–33.

Chimba, D., Sando, T., Kwigizile, V. (2010). Effect of bus size and operation to crash occurrences. *Accident Analysis & Prevention*, 42(6), 2063–2067.

Chin, H.C., Huang, H. (2009). Safety assessment of taxi drivers in Singapore. *Transportation Research Record: Journal of Transportation Research Board*, 2114(1), 47–56.

Chu, H.C. (2014). Assessing factors causing severe injuries in crashes of high-deck buses in long-distance driving on freeways. *Accident Analysis & Prevention*, 62, 130–136.

Chung, Y. (2018). Injury severity analysis in taxi-pedestrian crashes: An application of reconstructed crash data using a vehicle black box. *Accident Analysis & Prevention*, 111, 345–353.

Chung, Y. (2022). An application of in-vehicle recording technologies to analyze injury severity in crashes between taxis and two-wheelers. *Accident Analysis & Prevention*, 166, 106541.

Damsere-Derry, J., Adanu, E.K., Ojo, T.K., Sam, E.F. (2021). Injury-severity analysis of intercity bus crashes in Ghana: A random parameters multinomial logit with heterogeneity in means and variances approach. *Accident Analysis & Prevention*, 160, 106323.

Das, S., Jafari, M., Hossain, A., Chakraborty, R., Mimi, M.S. (2024). Toll road crash severity using mixed logit model incorporating heterogeneous mean structures.

Transportmetrica A: Transport Science, doi.org/10.1080/23249935.2024.2343755.

Feng, S., Li, Z., Ci, Y., Zhang, G. (2016). Risk factors affecting fatal bus accident severity: Their impact on different types of bus drivers. *Accident Analysis & Prevention*, 86, 29–39.

He, C., Xu, P., Pei, X., Wang, Q., Yue, Y., Han, C. (2024). Fatigue at the wheel: A non-visual approach to truck driver fatigue detection by multi-feature fusion. *Accident Analysis & Prevention*, 199, 107511.

Hosseinzadeh, A., Moeinaddini, A., Ghasemzadeh, A. (2021). Investigating factors affecting severity of large truck-involved crashes: Comparison of the SVM and random parameter logit model. *Journal of Safety Research*, 77, 151–160.

Huang, H., Siddiqui, C., Abdel-Aty, M. (2011). Indexing crash worthiness and crash aggressivity by vehicle type. *Accident Analysis & Prevention*, 43(4), 1364–1370.

Kaplan, S., Prato, C.G. (2012). Risk factors associated with bus accident severity in the United States: A generalized ordered logit model. *Journal of Safety Research*, 43(3), 171–180.

Lam, L.T. (2004). Environmental factors associated with crash-related mortality and injury among taxi drivers in New South Wales, Australia. *Accident Analysis & Prevention*, 36(5), 905–908.

Ma, Q., Yang, H., Xie, K., Wang, Z., Hu, X. (2019). Taxicab crashes modeling with informative spatial autocorrelation. *Accident Analysis & Prevention*, 131, 297–307.

Ma, J., Cao, Q., Ren, G., Yang, Y., Deng, Y., Li, J. (2024). Exploring the heterogeneous effects of riding behaviors and road conditions on delivery rider severities in scooter-style electric bicycle crashes involving vehicles. *International Journal of Injury Control and Safety Promotion*, 31(2), 165–180.

Mallia, L., Lazuras, L., Violani, C., Lucidi, F. (2015). Crash risk and aberrant driving behaviors among bus drivers: The role of personality and attitudes towards traffic safety. *Accident Analysis & Prevention*, 79, 145–151.

Mannering, F.L., Shankar, V., Bhat, C.R. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research*, 11, 1–16.

Meng, F., Xu, P., Wong, S.C., Huang, H., Li, Y.C. (2017). Occupant-level injury severity analyses for taxis in Hong Kong: A Bayesian space-time logistic model. *Accident Analysis & Prevention*, 108, 297–307.

Naik, B., Tung, L.W., Zhao, S., Khattak, A.J. (2016). Weather impacts on single-vehicle truck crash injury severity. *Journal of Safety Research*, 58, 57–65.

Nasri, M., Aghabayk, K. (2021). Assessing risk factors associated with urban transit bus involved accident severity: A case study of a Middle East country. *International Journal of Crashworthiness*, 26(4), 413–423.

Park, H.C., Kim, D.K., Kho, S.Y., Park, P.Y. (2017). Cross-classified multilevel models for severity of commercial motor vehicle crashes considering heterogeneity among companies and regions. *Accident Analysis & Prevention*, 106, 305–314.

Park, H.C., Joo, Y.J., Kho, S.Y., Kim, D.K., Park, B.J. (2019). Injury severity of bus–pedestrian crashes in South Korea considering the effects of regional and company factors. *Sustainability*, 11(11), 3169.

Peng, Z., Wang, Y., Truong, L.T. (2022). Individual and combined effects of working conditions, physical and mental conditions, and risky driving behaviors on taxi crashes in China. *Safety Science*, 151, 105759.

Peng, Z., Zuo, J., Ji, H., Yuan, R., Wang, Y. (2024). A comparative analysis of risk factors in taxi-related crashes using XGBoost and SHAP. *International Journal of Injury Control and Safety Promotion*, doi.org/10.1080/17457300.2024.2349555.

Prato, C.G., Kaplan, S. (2014). Bus accident severity and passenger injury: Evidence from Denmark. *European Transport Research Review*, 6, 17–30.

Rahman, M., Kattan, L., Tay, R. (2011). Injury risk in collisions involving buses in Alberta, Canada. *Transportation Research Record: Journal of Transportation Research Board*, 2265(1), 13–26.

Riccardi, M.R., Mauriello, F., Scarano, A., Montella, A. (2023). Analysis of contributory factors of fatal pedestrian crashes by mixed logit model and association rules. *International Journal of Injury Control and Safety Promotion*, 30(2), 195–209.

Saeed, T.U., Hall, T., Baroud, H., Volovski, M.J. (2019). Analyzing road crash frequencies with uncorrelated and correlated random-parameters count models: An empirical assessment of multilane highways. *Analytic Methods in Accident Research*, 23, 100101.

Sam, E.F., Daniels, S., Brijs, K., Brijs, T., Wets, G. (2018). Modelling public bus/minibus transport accident severity in Ghana. *Accident Analysis & Prevention*, 119, 114–121.

Samerei, S.A., Aghabayk, K., Mohammadi, A., Shiwakoti, N. (2021a). Data mining approach to model bus crash severity in Australia. *Journal of Safety Research*, 76, 73–82.

Samerei, S.A., Aghabayk, K., Shiwakoti, N., Karimi, S. (2021b). Modelling bus-pedestrian crash severity in the state of Victoria, Australia. *International Journal of Injury Control and Safety Promotion*, 28(2), 233–242.

Sarker, M.A.A., Rahimi, A., Azimi, G., Jin, X. (2023). Injury severity of single-vehicle large-truck crashes: Accounting for heterogeneity. *International Journal of Injury Control and Safety Promotion*, 30(4), 571–581.

Se, C., Champahom, T., Jomnonkwao, S., Ratanavaraha, V. (2022). Motorcyclist injury severity analysis: A comparison of artificial neural networks and random parameter model with heterogeneity in means and variances. *International Journal of Injury Control and Safety Promotion*, 29(4), 500–515.

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., Milton, J. (2017). Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research*, 15, 41–55.

Shen, J., Gu, H., Feng, X., Yu, M., Zheng, C. (2022). Investigation of factors contributing to bus-crash severity based on extended hierarchical ordered probit model with heteroscedasticity. *Journal of Transportation Safety & Security*, 14(9), 1531–1551.

Sun, Z., Wang, Z., Qi, X., Wang, D., Gu, X., Wang, J., Lu, H., Chen, Y. (2024). Understanding key contributing factors on the severity of traffic violations by elderly drivers: A hybrid approach of latent class analysis and XGBoost based

SHAP. *International of Injury Control and Safety Promotion*, 31(2), 273–293.

Tamakloe, R., Adanu, E.K., Atandzi, J., Das, S., Lord, D., Park, D. (2023). Stability of factors influencing walking-along-the-road pedestrian injury severity outcomes under different lighting conditions: A random parameters logit approach with heterogeneity in means and out-of-sample predictions. *Accident Analysis & Prevention*, 193, 107333.

Tamakloe, R., Hong, J., Park, D. (2020). A copula-based approach for jointly modeling crash severity and number of vehicles involved in express bus crashes on expressways considering temporal stability of data. *Accident Analysis & Prevention*, 146, 105736.

Tamakloe, R., Lim, S., Sam, E.F., Park, S.H., Park, D. (2021). Investigating factors affecting bus/minibus accident severity in a developing country for different subgroup datasets characterised by time, pavement, and light conditions. *Accident Analysis & Prevention*, 159, 106268.

Tang, J., Liang, J., Han, C., Li, Z., Huang, H. (2019). Crash injury severity analysis using a two-layer Stacking framework. *Accident Analysis & Prevention*, 122, 226–238.

Truong, L.T., Nguyen, H.T., Tay, R. (2020). A random parameter logistic model of fatigue-related motorcycle crash involvement in Hanoi, Vietnam. *Accident Analysis & Prevention*, 144, 105627.

Wang, C., Shao, Y., Ye, F., Zhu, T. (2024). Injury severity analysis of e-bike riders in China based on the in-vehicle recording video crash data: A random parameter ordered logit model. *International Journal of Injury Control and Safety Promotion*, doi.org/10.1080/17457300.2024.2385102.

Wang, J., Huang, H., Xu, P., Xie, S., Wong, S.C. (2020). Random parameter probit models to analyze pedestrian red-light violations and injury severity in pedestrian–motor vehicle crashes at signalized crossings. *Journal of Transportation Safety & Security*, 12(6), 818–837.

Wang, X., Jiao, Y., Huo, J., Li, R., Zhou, C., Pan, H., Chai, C. (2021). Analysis of safety climate and individual factors affecting bus drivers' crash involvement using a two-level logit model. *Accident Analysis & Prevention*, 154, 106087.

Wang, Y., Li, L., Feng, L., Peng, H. (2014). Professional drivers' views on risky driving behaviors and accident liability: A questionnaire survey in Xining, China. *Transportation Letters*, 6(3), 126–135.

Waseem, M., Ahmed, A., Saeed, T.U. (2019). Factors affecting motorcyclists' injury severities: An empirical assessment using random parameters logit model with heterogeneity in means and variances. *Accident Analysis & Prevention*, 123, 12–19.

Washington, S.P., Karlaftis, M.G., Mannering, F. (2020). *Statistical and Econometric Methods for Transportation Data Analysis (3rd Edition)*. Chapman and Hall/CRC.

Wu, Q., Zhang, G., Zhu, X., Liu, X., Tarefder, R. (2016). Analysis of driver injury severity in single-vehicle crashes on rural and urban roadways. *Accident Analysis & Prevention*, 94, 35–45.

Xiao, D., Yuan, Q., Xu, X., Zhang, S. (2024). Investigating injury severity interaction between urban and urban-rural fringe areas: A grouped random parameters seemingly unrelated bivariate probit approach. *International Journal of Injury Control and Safety Promotion*, 31(1), 30–37.

Xu, P., Bai, L., Pei, X., Wong, S.C., Zhou, H. (2022). Uncertainty matters: Bayesian modeling of bicycle crashes with incomplete exposure data. *Accident Analysis & Prevention*, 165, 106518.

Xu, P., Xie, S., Dong, N., Wong, S.C., Huang, H. (2019). Rethinking safety in numbers: Are intersections with more crossing pedestrians really safer? *Injury Prevention*, 25(1), 20–25.

Xu, X., Xie, S., Wong, S.C., Xu, P., Huang, H., Pei, X. (2016). Severity of pedestrian injuries due to traffic crashes at signalized intersections in Hong Kong: A Bayesian spatial logit model. *Journal of Advanced Transportation*, 50(8), 2015–2028.

Ye, Y., Zheng, P., Liang, H., Chen, X., Wong, S.C., Xu, P. (2024). Safety or efficiency? Estimating crossing motivations of intoxicated pedestrians by leveraging the inverse reinforcement learning. *Travel Behaviour and Society*, 35, 100760.

Yoon, S., Kho, S.Y., Kim, D.K. (2017). Effect of regional characteristics on injury severity in local bus crashes: Use of hierarchical ordered model. *Transportation*

Research Record: Journal of Transportation Research Board, 2647(1), 1–8.

Yu, M., Zheng, C., Ma, C. (2020). Analysis of injury severity of rear-end crashes in work zones: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research*, 27, 100126.

Yuan, R., Gan, J., Peng, Z., Xiang, Q. (2022). Injury severity analysis of two-vehicle crashes at unsignalized intersections using mixed logit models. *International Journal of Injury Control and Safety Promotion*, 29(3), 348–359.

Zeng, Q., Wen, H., Huang, H. (2016). The interactive effect on injury severity of driver-vehicle units in two-vehicle crashes. *Journal of Safety Research*, 59, 105–111.

Zeng, Q., Gu, W., Zhang, X., Wen, H., Lee, J., Hao, W. (2019). Analyzing freeway crash severity using a Bayesian spatial generalized ordered logit model with conditional autoregressive priors. *Accident Analysis & Prevention*, 127, 87–95.

Zeng, Q., Wang, Q., Zhang, K., Wong, S.C., Xu, P. (2023). Analysis of the injury severity of motor vehicle–pedestrian crashes at urban intersections using spatiotemporal logistic regression models. *Accident Analysis & Prevention*, 189, 107119.

Zhai, X., Huang, H., Sze, N. N., Song, Z., Hon, K.K. (2019). Diagnostic analysis of the effects of weather condition on pedestrian crash severity. *Accident Analysis & Prevention*, 122, 318–324.

Zhou, H., Yuan, C., Dong, N., Wong, S. C., Xu, P. (2020). Severity of passenger injuries on public buses: A comparative analysis of collision injuries and non-collision injuries. *Journal of Safety Research*, 74, 55–69.