

ORIGINAL PAPER

Open Access



Exploring injury severity of pedestrian-vehicle crashes at intersections: unbalanced panel mixed ordered probit model

Željko Šarić¹, Xuecai Xu², Daiquan Xiao^{2*} and Joso Vrkljan³

Abstract

Although the pedestrian deaths have been declining in recent years, the pedestrian-vehicle death rate in Croatia is still pretty high. This study intended to explore the injury severity of pedestrian-vehicle crashes with panel mixed ordered probit model and identify the influencing factors at intersections. To achieve this objective, the data were collected from Ministry of the Interior, Republic of Croatia from 2015 to 2018. Compared to the equivalent random-effects and random parameter ordered probit models, the proposed model showed better performance on goodness-of-fit, while capturing the impact of exogenous variables to vary among the intersections, as well as accommodating the heterogeneity issue due to unobserved effects. Results revealed that the proposed model can be considered as an alternative to deal with the heterogeneity issue and to decide the factor determinants. The results may provide beneficial insight for reducing the injury severity of pedestrian-vehicle crashes.

Keywords: Injury severity, Pedestrian-vehicle crash, Unbalanced panel mixed ordered probit model, Random parameter ordered probit model

1 Introduction

The latest traffic safety data provided by the National Highway Traffic Safety Administration (NHTSA, 2020) summarized that pedestrian deaths reduced from 8342 to 5320 in Europe, a declines of 36 percent from 2009 to 2018, while in American the number increased 51 percent from 4109 to 6227. Although the pedestrian deaths have been decreased apparently in Europe, the average death rate is still kept over 443 annually, and more attention should be paid to deal with the pedestrians, and even the vulnerable road users. Among all the pedestrian deaths, pedestrian-vehicle crashes account for most of the proportion in Europe, same as in Croatia. According to the statistics by Ministry of the Interior, Croatia from 2015 to 2018 there were totally 7155 pedestrian-vehicle

collisions, and about 2385 annually. Although the total number of death is not large (293 in total), the death rate is 58.6 per year, which is still relatively high.

In Croatia, not all the intersections are designed and labeled in a standard way, thus there still exists deficiency at some intersections, e.g. no pedestrian zebra crossing, no countdown light, inappropriate signal phase, bad lighting, etc., all of which may lead to the confliction between pedestrians and vehicles, and cause pedestrian injury. Consequently, in order to reduce the pedestrian-vehicle crashes and improve the pedestrian safety level, a variety of measures are being taken to find out the influencing factors and deal with the injury severity of pedestrian-vehicle crashes.

Therefore, the objective of this study includes the two main aspects: empirically to figure out the influencing factors for the injury severity of pedestrian-vehicle crashes at intersections; theoretically (1) to capture the impact of exogenous variables to vary among the intersections and (2) to accommodate the heterogeneity

*Correspondence: xiaodaiquan@hust.edu.cn

² School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
Full list of author information is available at the end of the article

issue due to unobserved effects. The results may provide potential insight for reducing the injury severity of pedestrian-vehicle crashes.

2 Literature review

As for the pedestrian-related crashes, currently three types of methods, such as on-site investigation, mathematical modeling, simulation, etc., have been utilized to assess the pedestrian injury severity, in which econometric modeling has displayed potential promising [30].

The primary method to address pedestrian injury severity began with the probit or logistic regression model. Zajac and Ivan [31] employed ordered probit model to investigate the factors influencing injury severity of pedestrian crashes in Connecticut. The results showed that clear roadway width, vehicle type, driver alcohol involvement, 65 years or older pedestrians and pedestrian alcohol involvement were significant factors. Similarly, Sze and Wong [26] adopted binary logistic regression model to evaluate the injury risk of pedestrian casualties. It was found that demographic, road environment, and other factors were controlled for pedestrian injury risk while temporal variation was considered initially. Focused on age and pedestrian injury severity, Kim et al. [11] conducted heteroskedastic logit analysis and the results displayed pedestrian age increased the probability of fatal pedestrian injury. Pointed at injury severity in taxi-pedestrian crashes, Chung [4] utilized ordered probit model and the results revealed that pedestrian legs were the first impact region with less severe in crashes. From the perspective of time series, Batouli et al. [1] analyzed the pedestrian-vehicle crash injury severity factors with logistic regression model. It was found that pedestrian age and pedestrian impairment were significantly concerned with the severity.

Panel data (random-effects, random-parameter or mixed) logit or probit models have been extended to accommodate the unobserved heterogeneity of pedestrian injury severity with cross-sectional and time-series data in recent years. Eluru et al. [8] developed mixed generalized ordered response logit model to examine pedestrian injury severity level. The results suggested that age, speed limit, location and time-of-day were significant influencing factors. Kim et al. [12] employed a mixed logit model to address unobserved heterogeneity in pedestrian injury severity. It was shown that heterogeneity in the mean of the random parameters for pedestrian-solely-at-fault collision indicator was associated with pedestrian gender while that for the traffic-sign and motorist-back-up indicators was associated with pedestrian age. Similar study performed by Li and Fan [16], Li et al. [17] verified mixed logit model. Extendedly, Xin et al. [28] quantified the

effects of neighborhood characteristics and built environment on pedestrian injury severity with random parameters generalized ordered probit model. For elderly pedestrian indicator the parameter was random with significant heterogeneity in both mean and variance and the elderly pedestrians involved in intersection-related crashes were suffered severe injury with high probability. As for the pedestrian red-light violations, Wang et al. [27] investigated the influencing factors of pedestrian-motor crashes at signalized crossings. Random parameter probit models were adopted to account for individual-specific heterogeneity, and pedestrian age of 12 to 25 years was likely to violate the red-light, as well as deep night, rainy weather, elderly pedestrians and bus crashes. From the perspective of hierarchical models, Kim et al. [13] established a hierarchical ordered model with crash features as lower-level variables and municipality features as upper-level. Elderly pedestrians were found to be significant factors of pedestrian injury and the proposed model explained a 7% unexplained variation in injury severity outcomes.

Combination with other approaches (e.g. clustering analysis, spatial and temporal analysis) has improved the pedestrian injury severity spatially and temporally. Mohamed et al. [19] combined data mining and regression methods to determine the influencing factors of pedestrian injury severity. A latent class with ordered probit model revealed the results greatly in New York whereas K-means with a multinomial logit model provided the results appropriately in Montreal. The findings showed that pedestrian age and other factors were significant factors of fatal crashes. Sasidharan et al. [23] employed the latent class cluster analysis to investigate pedestrian crash injury severities. The results suggested that this method was effective in reducing heterogeneity and revealing hidden relationships between crash severity levels and relevant factors. Identical method adopted by Zhao et al. [32] verified injury severity in pedestrian-train crashes at highway-rail grade crossings. From the perspective of space, Bhat et al. [2] proposed a spatial random coefficient multivariate count model to analyze pedestrian injury counts by severity level. Spatial heterogeneity, spatial dependency and spatial drift effects were accommodated and several groups of influencing factors were considered to be significant to reduce the number of pedestrian-vehicle crashes by severity level. By integrating spatial and temporal domain, Liu et al. [18] adopted geographically and temporally weighted ordinal logistic regression model to explore the correlates of pedestrian injury severity. It was found that pedestrian age, pedestrian position and other factors were highly related to pedestrian injury severity. Recent study by Song et al. [25] combined spatiotemporal analysis with hierarchical

Bayesian random-effects logit model to examine the factors of pedestrian-injury severities.

There have been a variety of other methods to analyze pedestrian crash injury severity. Li et al. [14] combined classification and regression tree (CART) with random forest to analyze pedestrian injury severity under different weather conditions. The results showed that elderly pedestrians and higher speed limit led to higher injury severities. Sasidharan and Menendez [22] applied partial proportional odds model to identify the influencing factors of pedestrian injury severities. Pedestrian age was found to significantly affect the injury severity levels. Xu et al. [30] investigated pedestrian injury severity and addressed the heterogeneity issue at signalized intersections. Bayesian quantile regression models were presented and the results revealed that age, pedestrian special circumstance and pedestrian contributory were likely to influence the injury severity. Extendedly, Mujalli et al. [20] applied Bayesian networks to determine impact factors of pedestrian-vehicle crash severity. The significant factors increasing the risk of severe injury or fatality were speed limit, lighting and weather condition.

To sum up, the evaluation on injury severity of pedestrian-vehicle crashes varies from pooled logistic or probit regression models, to panel data models, and to cluster analysis and spatial-temporal analysis, which addresses the injury severity from various perspectives. However, the current models or approaches mainly concentrate on one specific aspect. Therefore, the purpose of this study is to (1) capture the impact of exogenous variables to vary among the intersections and (2) accommodate the heterogeneity issue due to unobserved effects.

3 Methodology

Assume i ($i=1, 2, \dots, N$) to be an index to denote intersection, j ($j=1, 2, \dots, J$) be an index for different intersection i and k ($k=1, 2, 3, \dots, K$) to be an index to represent pedestrian injury severity. The latent propensity equation for injury severity at i th intersection and j th interval can be expressed as:

$$y_{ij}^* = (\alpha' + \delta'_i)z_{ij} + \varepsilon_i \tag{1}$$

where y_{ij}^* here denotes the pedestrian injury severity probability y_{ijk} by the thresholds ($\omega_0 = -\infty$ and $\omega_k = \infty$), z_{ij} represents an $(L \times 1)$ column vector of influencing factors on injury severity. α denotes a $(L \times 1)$ column vector of mean effects, and δ_i is another $(L \times 1)$ column vector of unobserved factors influencing z_i on the pedestrian injury severity probability for intersection i , and ε_i is random error term (i.i.d. across intersections i). Specifically, pedestrian injury severity y_{ij} is categorized no injuries, slight injury, severe injury and fatality, which matches the

attributes of ordered probit model. Due to the fixed- and random-effects considered with panel data framework, panel mixed ordered probit model can reflect the features of exogenous variables and heterogeneity issue.

Different from the conventional ordered probit model with maximum likelihood estimation, quasi-likelihood based method is employed to estimate the panel mixed ordered probit model. In Eq. (1), the α , ω and δ_i are required to be estimated, so it is assumed that

$$E(y_{ik}|z_{ik}) = H_{ik}(\alpha, \omega, \delta_i), 0 \leq H_{ik} \leq 1, \sum_{k=1}^K H_{ik} = 1 \tag{2}$$

where H_{ik} represents the ordered probit probability for injury severity category k , which is defined as: $P_{ijk} = \{G[\omega_k - \{(\alpha' + \delta'_i)z_i\}] - G[\omega_{k-1} - \{(\alpha' + \delta'_i)z_i\}]\}$ where $G(\cdot)$ denotes the cumulative distribution of the standard normal distribution.

Thus the quasi-likelihood function at intersection i for a given δ_i can be described as:

$$L_i(\alpha, \omega|\delta_i) = \prod_{j=1}^J \prod_{k=1}^K \{G[\omega_k - \{(\alpha' + \delta'_i)z_i\}] - G[\omega_{k-1} - \{(\alpha' + \delta'_i)z_i\}]\}^{d_{ik}} \tag{3}$$

where d_{ik} represents the proportion of injury severity category k .

The unconditional likelihood function for intersection i can be calculated as follows:

$$L_i(\alpha, \omega, \delta_i) = \int_{\delta_i} L_i(\alpha, \omega|\delta_i) dF(\delta_i) \tag{4}$$

where F denotes the multi-dimensional cumulative normal distribution. Therefore, the quasi log-likelihood function can be converted as:

$$L(\varphi) = \sum_i L_i(\alpha, \omega, \delta_i) \tag{5}$$

More details about the estimation can be referred to Eluru et al. [9] and Chen et al. [3].

4 Data description

The dataset was collected from Traffic Accident Database System maintained by the Ministry of the Interior, Republic of Croatia from 2015 to 2018. The population from Traffic Accident Database System covered the four main aspects: injury severity features, pedestrian status, vehicle characteristics, roadway features and environmental profiles.

Totally there are about 7155 injury severity selected, in which about 2260 occurred at intersections (shown in Fig. 1). Among the intersections selected, 1995 samples were determined for the analysis after eliminating some observations without data or incomplete. As stated above, in Croatia the pedestrian-vehicle injury severity levels are typically categorized as no injuries, slight injury, severe injury and fatality, which can be ranked as ordered level, thus dependent variables were considered as ordered probit model; By integrating the dataset from different counties with different years, panel data framework can be constructed. However, since the data are not the same at each county of every year, unbalanced panel data model should be developed; In order to address the fixed- and random-effects simultaneously, mixed effects were considered, therefore, the final unbalanced panel mixed ordered probit model was proposed.

Since pedestrians play significant role in pedestrian-vehicle injury severity, pedestrian status, such as gender, age, alcohol or not, number of participants in total, etc., and vehicle driver circumstances were collected to examine whether either part should take more responsibility,

while the vehicle-related variables include the total vehicles involved and vehicle types.

The roadway characteristics contain the intersection type (e.g. Y-intersection, T-intersection, four-intersection, roundabout, etc.), roadway conditions (e.g. dry, wet, ice, snow, etc.), speed limit and vertical/horizontal signalization, while the injury includes the day, location, severity level and the environment conditions collect the visibility conditions, public lighting, weather conditions, and state of environment.

In order to evaluate the proposed models, all the parameters are digitalized and listed in Table 1, including the dependent variable, categorical variables, continuous variables and indicator variables, respectively.

5 Results

Based on the variables selected from the four aspects, the correlation among independent variables needs to be examined before running the model. The Pearson correlation test was conducted to avoid the co-linearity. Shown from the test result, public lighting is highly related to visibility, number of participants in total is highly related to number of participants no injuries, and the number

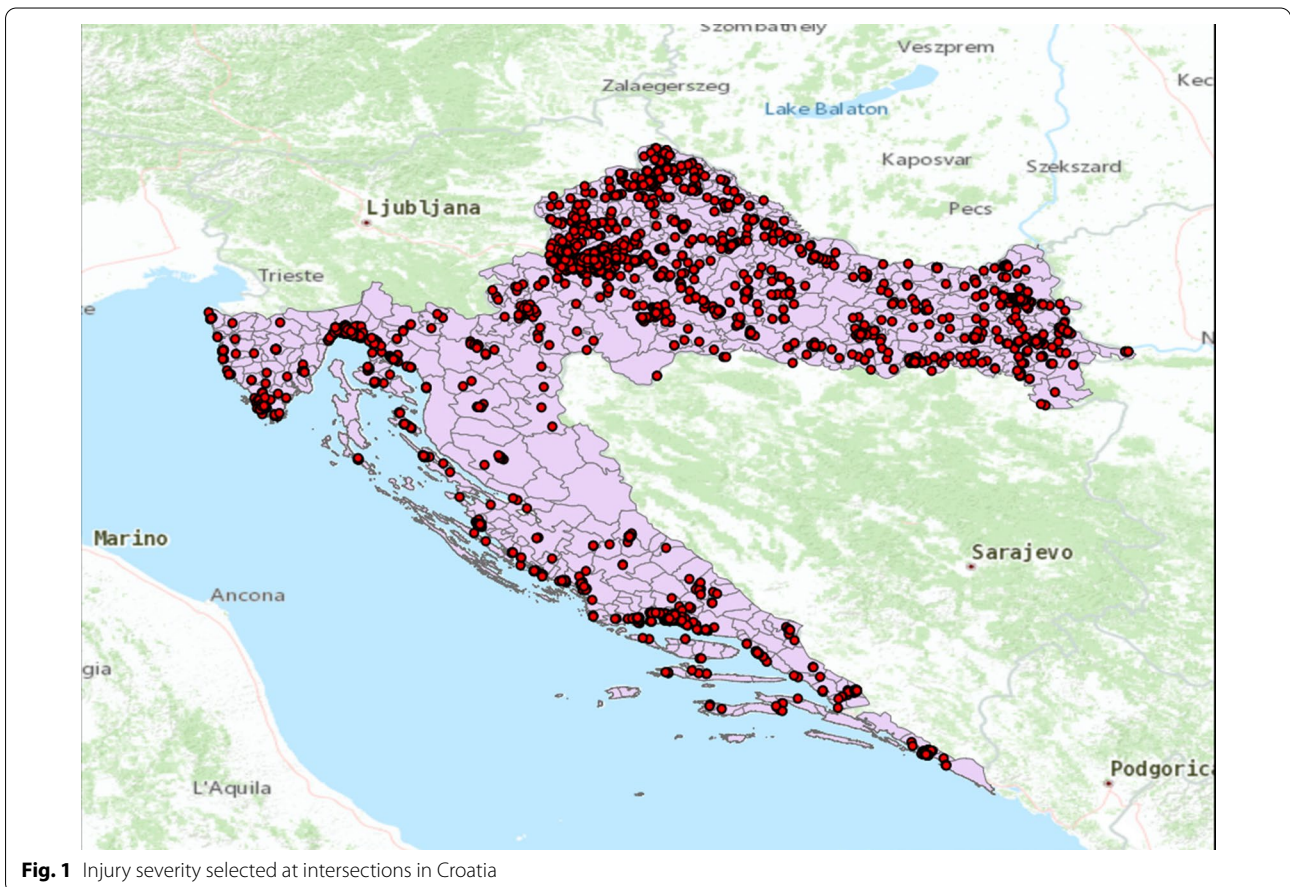


Fig. 1 Injury severity selected at intersections in Croatia

Table 1 Summary of parameter description

| Parameter | Category | Count (proportion) |
|----------------------------|--|--------------------|
| (i) Dependent variables | | |
| Year | 2015 | 503 (25.2%) |
| | 2016 | 520 (26.1%) |
| | 2017 | 517 (25.9%) |
| | 2018 | 455 (22.8%) |
| Injury severity | 1-No injuries | 68 (3.4%) |
| | 2-Slight injury | 1375 (68.9%) |
| | 3-Severe injury | 513 (25.7%) |
| | 4-Fatality | 39 (2.0%) |
| (ii) Categorical variables | | |
| Pedestrian age | 0-Less than 15 years old | 147 (7.4%) |
| | 1-15- 65 years old | 1245 (62.4%) |
| | 2-More than 65 years old | 603 (30.2%) |
| Driver circumstances | 1-Delayed detection of danger | 86 (4.3%) |
| | 2-Disregarding the advantage of passing | 146 (7.3%) |
| | 3-Disregarding the light sign | 77 (3.9%) |
| | 4-Disregarding the light sign—pedestrian | 128 (6.4%) |
| | 5-Driving at an insufficient distance | 5 (0.2%) |
| | 6-Improper detour | 9 (0.5%) |
| | 7-Improper inclusion in traffic | 25 (1.2%) |
| | 8-Improper driving(overtaking/reversing/passing/speed/turning) | 281 (14.1%) |
| | 9-Negligent handling of the vehicle | 16 (0.8%) |
| | 10-Non-use of a marked pedestrian crossing | 67 (3.4%) |
| | 11-Other driver mistake | 720 (36.1%) |
| | 12-Other pedestrian mistake | 97 (4.9%) |
| | 13-Speed inappropriate for conditions | 337 (16.9%) |
| | 14-The unexpected appearance of danger on the road | 1 (0.01%) |
| Vehicle type | 1-Bicycle | 28 (1.4%) |
| | 2-Personal automobile | 1428 (71.6%) |
| | 3-Bus/truck/tram | 132 (6.6%) |
| | 4-Motorcycle | 49 (2.5%) |
| Intersection type | 0-Unknown | 358 (17.9%) |
| | 0-Y-intersection | 118 (5.9%) |
| | 1-T-intersection | 919 (46.1%) |
| | 2-Four-intersection | 907 (45.5%) |
| Road conditions | 3-Roundabout | 51 (2.5%) |
| | 0-Dry | 1506 (75.5%) |
| | 1-Wet (including snow) | 482 (24.2%) |
| Speed limit | 2-Other | 7 (0.3%) |
| | 0 < 50 km/h | 168 (8.4%) |
| | 1 = 50 km/h | 1755 (88.0%) |
| Visibility conditions | 2 > 50 km/h | 72 (3.6%) |
| | 0-Day | 1368 (68.6%) |
| | 1-Dawn/twilight | 49 (2.4%) |
| Public lighting | 2-Night | 578 (29.0%) |
| | 0-In operation | 923 (46.2%) |
| | 1-Out of order | 1037 (51.9%) |
| | 2-No lighting | 35 (1.8%) |

Table 1 (continued)

| Parameter | Category | Count (proportion) | | | |
|---------------------------------------|---|--------------------|------|-----|-----|
| Weather conditions | 1-Clear(brightly, cloudy) | 1663 (83.3%) | | | |
| | 2-Rain/snow | 311 (15.6%) | | | |
| | 3-Fog | 21 (1.1%) | | | |
| Condition of vertical signalization | 0-Good | 1947 (97.6%) | | | |
| | 1-Bad/damaged | 6 (0.3%) | | | |
| | 2-No vertical signalization | 42 (2.1%) | | | |
| Condition of horizontal signalization | 0-Good | 1936 (97.0%) | | | |
| | 1-Bad/deficient | 13 (0.7%) | | | |
| | 2-No horizontal signalization | 46 (2.3%) | | | |
| | | Mean | SD | Min | Max |
| (iii) Continuous variables | | | | | |
| Total vehicle | Total vehicles involved | 1.02 | 0.16 | 1 | 4 |
| NoPT | No. of participants in total | 2.13 | 0.46 | 1 | 7 |
| NoPD | No. of participants dead | 0.02 | 0.14 | 0 | 2 |
| NoPS | No. of participants slightly injured | 0.26 | 0.45 | 0 | 2 |
| NoPI | No. of participants severely injured | 0.74 | 0.53 | 0 | 4 |
| NoPN | No. of participants no injuries | 1.10 | 0.45 | 0 | 6 |
| (iv) Indicator variables | | | | | |
| Day | 1 = Weekday, 0 = weekend | 0.83 | | 0 | 1 |
| Gender | 1 = Male, 0 = Female | 0.37 | | 0 | 1 |
| Alcohol | 1 = Yes, 0 = No | 0.03 | | 0 | 1 |
| Environment | 1 = Environment arranged, 0 = poorly or deficiency in environment | 0.98 | | | |

of participants slightly injured to seriously injured, thus, in the final results the six variables may not occur at the same time.

In order to make the comparison, random-effects, random parameter and the proposed ordered probit model are conducted to evaluate the injury severity. Table 2 provides the results of three models.

Shown from Table 2, several observations can be sorted out. First, the significant variables of random parameter and panel mixed ordered probit models are identical, but both models contain one more significant variable (e.g. pedestrian gender) than random-effects ordered probit model. Second, the AIC and BIC values from proposed model (2358.907 and 2402.489) are smaller than those from random-effects (2362.854 and 2418.838) and random parameter models (2365.307 and 2421.291) respectively, and the difference is beyond 5, which indicates the models are statistically different, while the values from random-effects model are very close to random parameter model. Third, although the quasi/log likelihood values at zero are the same, the proposed model converges at -1168.453 whereas the random-effects model and random parameter model converge at -1171.427 and -1172.653 correspondingly. Therefore,

the goodness-of-fit of the proposed model performs better, thus the following explanation would concentrate on the unbalanced panel mixed ordered probit model.

The first significant variable day is negatively related to pedestrian-vehicle injury severity, indicating that from weekend to weekday the injury severity level is decreasing from severe to no injuries. This implies that although there exist more injuries on weekdays, the severity level on weekends is more severe since more pedestrians go out on weekends. The finding is consistent with Dai [5], and Senserrick et al. [24].

In Table 2, pedestrian age, especially more than 65 years old, and gender are two significant variables influencing pedestrian-vehicle injury severity levels. Compared to pedestrians less than 65 years, the older pedestrians over 65 years old reveal a positive relation with injury severity, meaning that the older the pedestrians are, the more severe the injury is. Similar results were found by Eluru et al. [8] and Li and Fan [15]. The reason is such that it takes longer walking time for the elderly since their mobility is becoming slower than the younger, thus they are easily injury-prone and the severity is more severe; Pedestrian gender is negatively associated with injury severity of pedestrian-vehicle

Table 2 Results for random-effects, random parameter and unbalanced panel mixed ordered probit model

| Variable | Random-effects ordered probit | | | Random parameter ordered probit | | | Panel mixed ordered probit | | |
|-------------------------------------|-------------------------------|-------|--------|---------------------------------|-------|--------|----------------------------|-------|--------|
| | Coef | SE | z | Coef | SE | z | Coef | SE | z |
| Day | -0.198* | 0.078 | -2.55 | -0.214* | 0.077 | -2.77 | -0.203* | 0.078 | -2.62 |
| <i>Pedestrian age</i> | | | | | | | | | |
| > 65 years old | 0.367* | 0.065 | 5.58 | 0.375* | 0.065 | 5.75 | 0.364* | 0.066 | 5.54 |
| <i>Pedestrian gender</i> | | | | | | | | | |
| | | | | -0.152* | 0.061 | -2.50 | -0.149* | 0.061 | -2.43 |
| <i>Speed limit</i> | | | | | | | | | |
| > 50 km/h | 0.517* | 0.151 | 3.40 | 0.524* | 0.151 | 3.46 | 0.551* | 0.153 | 3.61 |
| <i>Vehicle type</i> | | | | | | | | | |
| | 0.124* | 0.032 | 3.79 | 0.111* | 0.032 | 3.45 | 0.127* | 0.033 | 3.87 |
| NoPT | 0.406* | 0.062 | 6.53 | 0.406* | 0.061 | 6.59 | 0.405* | 0.062 | 6.51 |
| NoPS | -1.498* | 0.059 | -25.37 | -1.498* | 0.058 | -25.54 | -1.506* | 0.059 | -25.43 |
| σ_{ϵ}^2 | 0.011 | 0.008 | | 0.011 | 0.008 | | 0.010 | 0.008 | |
| <i>Goodness-of-fit</i> | | | | | | | | | |
| No. of observations | 1995 | | | 1995 | | | 1995 | | |
| Quasi/Log likelihood at zero | -1591.702 | | | -1591.702 | | | -1591.702 | | |
| Quasi/Log Likelihood at convergence | -1171.427 | | | -1172.653 | | | -1168.453 | | |
| AIC | 2362.854 | | | 2365.307 | | | 2358.907 | | |
| BIC | 2418.838 | | | 2421.291 | | | 2402.489 | | |

Coef, coefficient; SE, standard error

*Significance at 95% confidence interval

crashes, which indicates that from female to male pedestrians, the injury severity levels are reduced. This implies that female pedestrians experience more severe injury, and the probability increases 14.9% if one more male pedestrian is converted into female.

Speed limit, especially over 50 km/h has a positive relation with injury severity. Compared to speed limit no more than 50 km/h, the probability of injury severity is increased from no injuries to severe ones, and previous studies [10, 21, 29] have confirmed this point.

Vehicle type plays a positive significant role in the injury severity of pedestrian-vehicle crashes, indicating that the probability of injury severity is increased with the vehicle type from non-motorized bicycles to motorcycles, which makes sense. Usually the injury severity by motorcycles is more severe than general cars because of the shocking force and irregular driving mobility, and some previous studies [6, 15] have verified this.

The last two significant variables, number of participants (total and slightly injured), are positively and negatively associated with injury severity of pedestrian-vehicle crashes, respectively. The positive relation indicates that the more participants involved the more injury severity the pedestrian-vehicle crashes, which is understandable in real life. On the contrary, the more participants with slightly injured, the less injury severity, implying that when more slightly injured participants are included, the injury is not much severe.

As stated above, there have been various methods and approaches about the injury severity analysis of pedestrian-vehicle crashes. However, most of the studies concentrate on the hotspots or influencing factors generally and individually, and there may exist heterogeneity issue due to unobserved factors. In this study, in order to deal with this, the panel mixed ordered probit model is proposed, which can capture the impact of exogenous variables to vary among the intersections and accommodate the heterogeneity issue due to unobserved effects.

Shown from Table 2, the closer examination of the comparison results provides some similarities and differences among the three models. First, the similarity lies in that for all the significant variables, random parameter model and the proposed model give the same factors. This indicates that both models are very identical. Secondly, the difference is that significant variables in random-effects model are fewer than the other two models. This implies that random-effects ordered probit model accommodates the heterogeneity issue partially. Consequently, the impact of exogenous variables and the heterogeneity issue for injury severity of pedestrian-vehicle crashes can be determined with panel mixed ordered probit model, which provides a potential insight of injury severity over current studies.

6 Conclusions

In this study, in order to capture the impact of exogenous variables to vary among the intersections and accommodate the heterogeneity issue due to unobserved effects, panel mixed ordered probit model was proposed to identify the significant influencing factors of injury severity for pedestrian-vehicle crashes at intersections. The results revealed that day, pedestrian gender and age, especially over 65 years old, speed limit over 50 km/h, vehicle type and number of participants were significant factors of injury severity.

Two main findings can be obtained from the results of the work. First, panel mixed ordered probit model emphasizes not only the mixed effects of errors as mixed ordered probit model, but expands the exogenous variables with mixed effects, which can capture the impact of exogenous variables to vary among the intersections and accommodate the heterogeneity issue. This is to our knowledge the first attempt to present this model in injury severity analysis of pedestrian-vehicle crashes and expands the range of injury severity analysis. Second, compared to random parameter and random effects models, the proposed model shows better performance from the goodness-of-fit, which can be considered as an alternative to deal with the heterogeneity issue completely.

Empirically, since more pedestrians go out on weekends and injury severity is more severe, corresponding measures should be taken to alert pedestrians at intersections or to prevent the vehicles from confliction with pedestrians, such as traffic lights countdown timer with sound, virtual pedestrian walk screen wall, etc.; At certain intersections with more elders over 65 years old and females, the walking time should be long enough to allow the older pedestrians walk through the intersections, and more attention should be paid to female pedestrians; Speed limit should strictly follow the roadway classification, and at intersections with more pedestrians the speed limit should be set up no more than 50 km/h; Motorcycles at certain areas with more pedestrians should not be permitted, e.g. leisure space, CBD area, because pedestrians are more vulnerable at those locations; At last, more education and lessons should be taken by drivers and pedestrians so that the number of total participants can be reduced.

Some drawbacks may need to be strengthened in the further study. More variables related to drivers and pedestrians status (e.g. pedestrian crossing behavior), vehicle conditions, roadway characteristics (e.g. signal phase), environment (e.g. residential area, school zone), and land use types around intersections Elias and Shifatan [7] so that injury severity can be reflected completely. Another issue is that the results of the work are founded

on the dataset from Croatia, and it is worthy of employing different data sources to ascertain the findings and transferability in the future. Future study may consider the injury severity spatially and temporally, so that spatial and temporal issues can be addressed clearly.

Acknowledgements

The authors are grateful for the data collection from Ministry of the Interior, Republic of Croatia.

Authors' contributions

Dr. ŽŠ provides the dataset and the topic idea; Dr. JV helps to process the data; Dr. XX writes the draft, and Dr. DX makes the revision and takes care of the submission. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and material

The data were provided by Ministry of the Interior, Republic of Croatia, and more requests about data can be referred to Dr. Željko Šarić (Email: zsarić@fpz.unizg.hr) at University of Zagreb, Croatia.

Declaration

Competing interests

The authors declare that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Author details

¹Department of Traffic Accident Expertise, Faculty of Transport and Traffic Sciences, University of Zagreb, Zagreb, Croatia. ²School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China. ³Polytechnic "Nikola Tesla", Bana Ivana Karlovića 16, 53000 Gospić, Croatia.

Received: 16 August 2021 Accepted: 23 November 2021

Published online: 11 December 2021

References

1. Batouli, G., Guo, M., Janson, B., & Marshall, W. (2020). Analysis of pedestrian-vehicle crash injury severity factors in Colorado 2006–2016. *Accident Analysis and Prevention*, *148*, 105782.
2. Bhat, C. R., Astroza, S., & Lavieri, P. S. (2017). A new spatial and flexible multivariate random-coefficients model for the analysis of pedestrian injury counts by severity level. *Analytic Methods in Accident Research*, *16*, 1–22.
3. Chen, F., Chen, S., & Ma, X. (2018). Analysis of hourly crash likelihood using unbalanced panel data mixed logit model and real-time driving environmental big data. *Journal of Safety Research*, *65*, 153–159.
4. Chung, Y. (2018). Injury severity analysis in taxi-pedestrian crashes: An application of reconstructed crash data using a vehicle black box. *Accident Analysis and Prevention*, *111*, 345–353.
5. Dai, D. (2012). Identifying clusters and risk factors of injuries in pedestrian vehicle crashes in a GIS environment. *Journal of Transport Geography*, *24*, 206–214.
6. Das, S. (2020). Identifying key patterns in motorcycle crashes: Findings from taxicab correspondence analysis. *Transportmetrica A-Transport Science*, *17*(4), 593–614.
7. Elias, W., & Shifatan, Y. (2014). Analyzing and modeling risk exposure of pedestrian children to involvement in car crashes. *Accident Analysis and Prevention*, *62*, 397–405.
8. Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention*, *40*(3), 1033–1054.
9. Eluru, N., Chakour, V., Chamberlain, M., & Miranda-Moreno, L. F. (2013). Modeling vehicle operating speed on urban roads in Montreal: A panel

- mixed ordered probit fractional split model. *Accident Analysis and Prevention*, 59, 125–134.
10. Haleem, K., Alluri, P., & Gan, A. (2015). Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis and Prevention*, 81, 14–23.
 11. Kim, J., Ulfarsson, G. F., Shankar, V. N., & Kim, S. (2008). Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic logit analysis. *Accident Analysis and Prevention*, 40(5), 1695–1702.
 12. Kim, J., Ulfarsson, G. F., Shankar, V. N., & Mannering, F. (2010). A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis and Prevention*, 42(6), 1751–1758.
 13. Kim, M., Kho, S., & Kim, D. (2017). Hierarchical ordered model for injury severity of pedestrian crashes in South Korea. *Journal of Safety Research*, 61, 33–40.
 14. Li, D., Ranjitkar, P., Zhao, Y., Yi, H., & Rashidi, S. (2017). Analyzing pedestrian crash injury severity under different weather conditions. *Traffic Injury Prevention*, 18(4), 427–430.
 15. Li, Y., & Fan, W. (2019). Modeling severity of pedestrian-injury in pedestrian-vehicle crashes with latent class clustering and partial proportional odds model: A case study of North Carolina. *Accident Analysis and Prevention*, 131, 284–296.
 16. Li, Y., & Fan, W. (2020). Mixed logit approach to modeling the severity of pedestrian-injury in pedestrian-vehicle crashes in North Carolina: Accounting for unobserved heterogeneity. *Journal of Transportation Safety & Security*. <https://doi.org/10.1080/19439962.2020.1821850>
 17. Li, Y., Song, L., & Fan, W. (2021). Day-of-the-week variations and temporal instability of factors influencing pedestrian injury severity in pedestrian-vehicle crashes: A random parameters logit approach with heterogeneity in means and variances. *Analytic Methods in Accident Research*, 29, 100152.
 18. Liu, J., Hainen, A., Li, X., Nie, Q., & Nambisan, S. (2019). Pedestrian injury severity in motor vehicle crashes: An integrated spatio-temporal modeling approach. *Accident Analysis and Prevention*, 132, 105272.
 19. Mohamed, M. G., Saunier, N., Miranda-Moreno, L. F., & Ukkusuri, S. V. (2013). A clustering regression approach: A comprehensive injury severity analysis of pedestrian-vehicle crashes in New York, US and Montreal, Canada. *Safety Science*, 54, 27–37.
 20. Mujalli, R. O., Garach, L., Lopez, G., & Al-Rousan, T. (2019). Evaluation of injury severity for pedestrian-vehicle crashes in Jordan using extracted rules. *Journal of Transportation Engineering Part A-Systems*, 145(7), 04019028.
 21. Prato, C. G., Kaplan, S., Patrier, A., & Rasmussen, T. K. (2018). Considering built environment and spatial correlation in modeling pedestrian injury severity. *Traffic Injury Prevention*, 19(1), 88–93.
 22. Sasidharan, L., & Menendez, M. (2019). Application of partial proportional odds model for analyzing pedestrian crash injury severities in Switzerland. *Journal of Transportation Safety & Security*, 11(1), 58–78.
 23. Sasidharan, L., Wu, K., & Menendez, M. (2015). Exploring the application of latent class cluster analysis for investigating pedestrian crash injury severities in Switzerland. *Accident Analysis and Prevention*, 85, 219–228.
 24. Senserrick, T., Boufous, S., DeRome, L., Ivers, R., & Stevenson, M. (2014). Detailed analysis of pedestrian casualty collisions in Victoria. *Australia. Traffic Injury Prevention*, 15(1), S197–S205.
 25. Song, L., Li, Y., Fan, W., & Wu, P. (2020). Modeling pedestrian-injury severities in pedestrian-vehicle crashes considering spatiotemporal patterns: Insights from different hierarchical Bayesian random-effects models. *Analytic Methods in Accident Research*, 28, 100137.
 26. Sze, N. N., & Wong, S. C. (2007). Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes. *Accident Analysis and Prevention*, 39(6), 1267–1278.
 27. Wang, J., Huang, H., Xu, P., Xie, S., & Wong, S. C. (2020). Random parameter probit model 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.
 28. Xin, C., Guo, R., Wang, Z., Lu, Q., & Lin, P. (2017). The effects of neighborhood characteristics and the built environment on pedestrian injury severity: A random parameters generalized ordered probability model with heterogeneity in means and variances. *Analytic Methods in Accident Research*, 16, 117–132.
 29. Xu, X., Luo, X., Ma, C., & Xiao, D. (2020). Spatial-temporal analysis of pedestrian injury severity with geographically and temporally weighted regression model in Hong Kong. *Transportation Research Part F-Traffic Psychology and Behavior*, 69, 286–300.
 30. Xu, X., Luo, X., Xiao, D., & Wong, S. C. (2019). Comparative analysis of Bayesian quantile regression models for pedestrian injury severity at signalized intersections. *Journal of Transportation Safety & Security*. <https://doi.org/10.1080/19439962.2019.1703867>
 31. Zajac, S. S., & Ivan, J. N. (2003). Factors influencing injury severity of motor vehicle-crossing pedestrian crashes in rural Connecticut. *Accident Analysis and Prevention*, 35(3), 369–379.
 32. Zhao, S., Iranitalab, A., & Khattak, A. J. (2019). A clustering approach to injury severity in pedestrian-train crashes at highway-rail grade crossings. *Journal of Transportation Safety & Security*, 11(3), 305–322.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)