Multimodal Transportation 3 (2024) 100109

Contents lists available at ScienceDirect

Multimodal Transportation

journal homepage: www.elsevier.com/locate/multra

Full Length Article

Injury severity analysis of highway-rail grade crossing crashes in non-divided two-way traffic scenarios: A random parameters logit model

Qiaoqiao Ren, Min Xu*

Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China

ARTICLE INFO

Keywords:
Highway-rail grade crossing crashes
Non-divided two-way traffic
Injury severity
Random parameters logit model
Unobserved heterogeneity

ABSTRACT

Highway-rail grade crossing (HRGC) crashes in non-divided two-way traffic scenarios have caused numerous fatalities and injuries over the years. Although crucial to the safety of multimodal transportation systems, these crossings have received little attention and previous studies did not fully account for the unobserved heterogeneity and its potential interactive effects. To bridge these gaps, the HRGC crashes occurring between 2019 and 2020 in the United States were collected from the Federal Railroad Administration's Office of Safety Analysis System. A random parameters logit model with heterogeneity in means was developed to investigate the impact of multiple factors associated with crossings, crashes, drivers, vehicles, and the environment. The present study indicates that did not stop behavior generates the random parameter with heterogeneity in means that is influenced by the dark and land with commercial power indicators. Furthermore, the findings show that factors such as estimated vehicle speed > 25 MPH, train speed > 45 MPH, going around the gate, old driver, female driver, motorcycle, and the driver was in vehicle indicators would increase the likelihood of more severe injury outcomes in HRGC crashes. Notably, the adverse crossing surface and truck indicators demonstrate unexpected marginal effects by reducing the likelihood of severe injury outcomes at non-divided two-way traffic HRGCs. This study emphasizes the importance of considering unobserved heterogeneity in the context of HRGC crashes. The findings can serve as a foundation for developing targeted interventions aimed at enhancing road and railway safety.

1. Introduction

Highway-rail grade crossing (HRGC) crashes are a significant public safety concern as they are responsible for numerous fatalities and injuries each year. According to the Federal Railroad Administration, in 2021, there were 2,146 accidents at HRGCs, resulting in 122 fatalities and 419 injuries in the United States (Federal Railroad Administration, 2021). Highway–rail grade crossings (HRGCs) are intersections where roadways cross railroads at the same level, and the interaction between non-divided two-way traffic highways and railways can be particularly dangerous. These types of road entities are characterized by their lack of a physical barrier between opposing lanes of traffic. The Federal Railroad Administration reports that around 95% of HRGC accidents occurred at non-divided two-way traffic HRGCs, indicating the necessity to further analyze accidents at these crossings (Federal Railroad Administration, 2021).

* Corresponding author.
E-mail addresses: qiaoqiao.ren@connect.polyu.hk (Q. Ren), min.m.xu@polyu.edu.hk (M. Xu).

https://doi.org/10.1016/j.multra.2023.100109
Received 5 April 2023; Received in revised form 12 June 2023; Accepted 8 July 2023

Despite being critical for the smooth operation and safety of multimodal transportation systems, non-divided two-way traffic HRGCs are rarely explored specifically in previous studies. Most of them focused on the crashes that occurred on the non-divided two-way highway and highway crossings only. For instance, Lombardi et al. (2017) investigated the factors contributing to age-related risks and involvement rates in fatal vehicle crashes that occurred at intersections in the United States. The findings revealed that a significant proportion of fatal crashes occurred at four-way intersections and non-divided two-way traffic ways were equally distributed between daylight and night-time conditions. Liu and Fan (2021) proposed a partial proportional odds model for examining the factors contributing to the severity of head-on crashes. Their analysis also revealed that head-on collisions occurring on two-way non-divided roadways were associated with more severe injuries than those on one-way roads. In a recent study conducted by Mahmud et al. (2022), the authors examined the safety risks associated with overtaking maneuvers on a bi-directional undivided two-lane highway in a heterogeneous traffic environment. The findings revealed that the speed difference between the overtaking and overtaken vehicles was a significant factor influencing the probability of severe conflicts. As indicated by Calvi et al. (2023), different types of median separation on two-lane rural roads can significantly affect driving behaviour (lateral positions and trajectories) and associated risks. Moreover, crashes are found to more commonly occur on non-divided two-way roads compared to other road types due to mixed traffic environments (Mahmud et al., 2022). Combining all road types (i.e., two-way traffic, divided traffic, and one-way traffic) when analyzing the injury severity of HRGC crashes can result in erroneous conclusions and improper countermeasures. Therefore, it is necessary to conduct separate studies focusing at non-divided two-way traffic HRGCs and investigate the factors affecting crashes in the context of this predominant traffic type.

2. Literature review

Regarding the safety of HRGCs, extensive studies have been conducted on crash frequency analysis, injury severity modeling, crash risk prediction, and risk mitigation measures evaluation (Gao et al., 2021; Kutela et al., 2022; Liu et al., 2022; Yan et al., 2023). A range of analytical techniques using statistical and data-driven approaches have been utilized to identify factors contributing to crashes and make predictions based on aggregated crash records or their subsets. For example, Hao et al. (2015) utilized an ordered probit model to explore the effect of age and gender on motor vehicle driver injury severity at HRGCs, and the analysis found that there were noteworthy variations in behavioral and physical characteristics between male and female drivers who were involved in accidents. In a comparative study conducted by Ghomi et al. (2016), the severity of vulnerable road users involved in HRGC crashes was assessed using three different methods, namely, an ordered probit model, association rules, and tree-based methods. The study results revealed that train speed, older road users, nighttime conditions, and female drivers were more likely to be associated with severe crashes. Mathew and Benekohal (2021) adopted the Zero Inflated Negative Binomial model and Empirical Bayes method for each type of warning device to predict HRGC crashes. The predicted values demonstrated better agreement with the field data in comparison to the newly released Federal Railroad Administration model. Based on the 19-year HRGC crash data in North Dakota, Gao et al. (2021) used the resampling approach to address the imbalanced data issue and employed several data-driven approaches to predict HRGC accidents. This study suggested that resample led to a considerable improvement in recall rate and the proposed deep learning approach exhibits superior predictive performance compared to other machine learning methods. Keramati et al. (2020) investigated the public HRGC crash severity in North Dakota from 1990 to 2018 by applying the random survival forest model. Their results indicated that installing audible devices to crossing with gates and standard flashing lights could result in a significant reduction of crash likelihood, PDO, injury, and fatal crashes by 49%, 52%, 46%, and 50%, respectively.

Numerous contributing factors to HRGC crashes have been studied, including crossing characteristics, traffic characteristics, vehicle type, driver behavior, pedestrian activity, weather, and visibility (Jamal et al., 2021; Mathew and Benekohal, 2021; Soleimani et al., 2019; Wu et al., 2022). For instance, Hao and Daniel (2013) utilized a conventional ordered probit model to conduct an analysis from1997 to 2006 and found that adverse weather conditions, vehicle speed exceeding 50 miles per hour (mph), and train speed surpassing 50 mph at the time of the collision were all associated with an increased probability of injury or fatality. In a subsequent investigation carried out by Hao and Daniel (2014), it was discovered that higher levels of Annual Average Daily Traffic (AADT) exceeding 10,000 vehicles per day were positively correlated with an elevated risk of injury and fatality at highway-rail grade crossings. Tjahjono et al. (2019) employed a conventional ordered logit modeling framework to identify determinant variables of the injury severity crashes at road-railway level crossings. The analysis outcomes revealed a significant correlation between fatal crashes and factors such as male drivers, rainy weather, and low traffic volume conditions.

While these studies have been valuable in identifying factors that contribute to the severity of HRGC crashes, they typically assumed that the coefficients were fixed for all observations. As a matter of fact, individuals' physical abilities, risk perceptions, and reactions to external stimuli differ, resulting in significant variations in the severity of injuries sustained in HRGC crashes. Meanwhile, it is important to acknowledge that while attempts are made to ensure the update and comprehensiveness of HRGC crash data, the collection of all pertinent information that could impact the occurrence of crashes is often limited by constraints related to costs and resources. In the case of HRGCs, several crash-specific features, including the traffic density, average speed, and unexpected pedestrian or animal trespassing may pose challenges to recording. Similarly, collecting human factors such as distraction, fatigue, the level of anxiety during the pandemic, or impairment of the driver may also bring considerable difficulties. This may unavoidably cause the unobserved heterogeneity (Mannering, 2018), and failing to account for unobserved variables or treating explanatory variables as fixed can lead to model misspecification, biased parameter estimates, and incorrect inferences (Alnawmasi and Mannering, 2019; Pervez et al., 2022).

Only a few studies have sought to address the methodological challenges arising from the unobserved heterogeneity in the domain of HRGC crash-related analysis. For example, Kutela et al. (2022) utilized the HRGC crash data spanning from 1999 to 2018 to

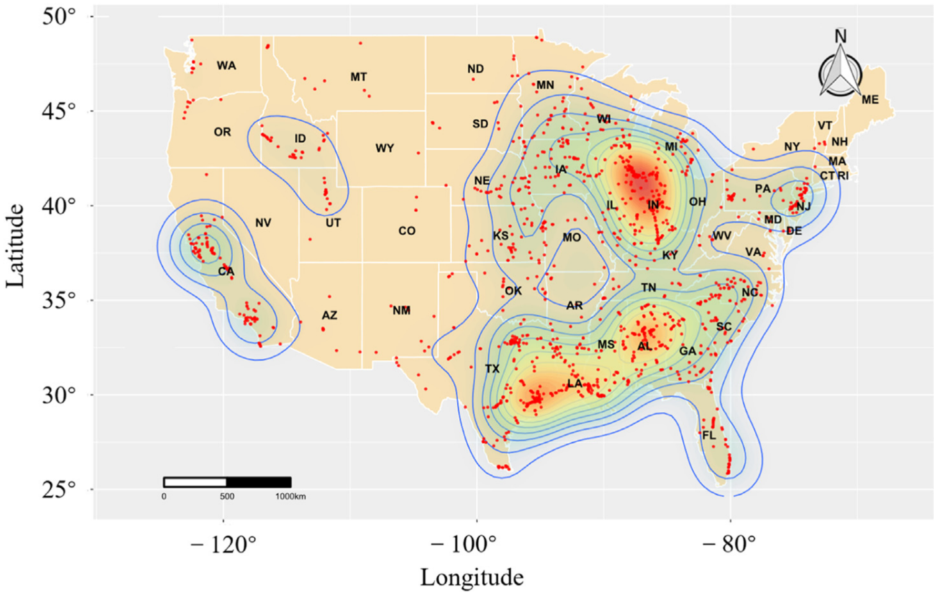

Fig. 1. The spatial distribution of highway–rail grade crossings crashes.

assess pre-crash driver behaviors using a mixed multinomial logit model. The study demonstrated a notable intra-class correlation coefficient, underscoring the need to integrate random-effect parameters into the model. Furthermore, the study unveiled that male drivers exhibit a tendency to ignore crossing gates, while older drivers tend to halt briefly and proceed before the train has entirely cleared the crossing. Ahmed et al. (2023) utilized the random parameters multinomial logit models with heterogeneity in the means and variances to account for the unobserved heterogeneity and spatial instability in the analysis of crash injury severity at HRGC using the crash data from Texas and California. The results indicated vehicle, train, driver, weather, and crossing characteristics that could significantly affect injury severity outcomes in state-specific models.

Given the unobserved heterogeneity arising from various factors has not been fully considered in the context of non-divided two-way traffic HRGCs, this study aims to establish an explicit analysis framework for identifying the underlying factors with the consideration of the unobserved heterogeneity. The structure of this paper is as follows. Section 3 presents the empirical setting and provides a descriptive statistical analysis of the explanatory variables. Section 4 introduces the method adopted in this study. Section 5 presents and discusses the model estimation results in detail. Finally, Section 6 concludes the study by summarizing the key findings, discussing policy implications, and highlighting directions for future research.

3. Data description

This study utilizes the data collected from the Federal Railroad Administration's (FRA) Office of Safety Analysis System. Records of crashes that occurred at non-divided two-way traffic HRGCs in the United States from 2020 to 2021 are obtained (Federal Railroad Administration, 2021). The research data is derived from two sources: (1) the FRA's Office of HRGC accident data and (2) the FRA's Office of current crossing inventory data. The accident database contains comprehensive information on factors such as time, weather, visibility, vehicle, and demographic characteristics of highway users, while the current highway-rail crossing inventory database provides details on each crossing's location, type, illumination, signals, and traffic conditions.

In this study, injury severity outcomes are coded using a three-point ordinal scale, with options for 1-Uninjured, 2-Injured, or 3-Killed. This discrete variable is designated as the dependent variable for the subsequent econometric analysis. The information in the current crossing inventory database is matched to the extracted crash database using the corresponding identification number assigned to each crossing to comprehensively identify factors influencing the crash injury severity. Furthermore, the integrated data is thoroughly verified and cleaned up to remove any crash observations with missing data in variables. The key determinants retrieved from the augmented database are then classified into crossing characteristics, crash characteristics, driver characteristics, vehicle characteristics, and environmental characteristics. Table 1 shows the detailed classification and descriptive statistics of independent variables. The final dataset used for analysis consists of 1503 crash records from 2020 to 2021, in which 1,009 individuals were uninjured (67.13%), 384 individuals were injured (25.55%), and 110 individuals were killed (7.32%). Fig. 1 presents the spatial distribution of HRGC crashes that occurred during the study period.

Table 1
Descriptive statistics of independent variables.

Category	Variables	Mean	Std. Dev.
Crossing characteristics	Private crossing type	0.020	0.144
	Public crossing type	0.980	0.144
	Crossing without signs or signals	0.020	0.142
	Crossing with signs or signals	0.980	0.142
	Unpaved Highway	0.100	0.298
	Paved Highway	0.900	0.298
	Land without commercial power	0.100	0.294
	Land with commercial power	0.900	0.294
	Highway speed limit <= 25 MPH	0.270	0.444
	Highway speed limit > 25 MPH	0.730	0.444
	Annual average daily traffic (AADT) <= 5,000	0.700	0.460
	Annual average daily traffic (AADT) > 5,000	0.300	0.460
	Estimated percent of trucks <= 10%	0.730	0.446
	Estimated percent of trucks > 10%	0.270	0.446
	Industry track	0.050	0.227
	Main track	0.880	0.321
	Siding track	0.010	0.077
	Yard track	0.060	0.230
	Both-side crossing warning	0.950	0.225
	Single-side crossing warning	0.050	0.225
	Crossing without illumination	0.620	0.486
	Crossing with illumination	0.380	0.486
	Dry crossing surface	0.860	0.349
	Adverse crossing surface*	0.140	0.349
Crash characteristics	Estimated vehicle speed <= 25 MPH	0.910	0.280
	Estimated vehicle speed > 25 MPH	0.090	0.280
	Train speed <= 45 MPH	0.840	0.371
	Train speed > 45 MPH	0.160	0.371
	Unobstructed view	0.960	0.194
	Obstructed view*	0.040	0.194
	Did not stop	0.380	0.486
	Stopped and then proceeded	0.060	0.239
	Stopped on the crossing	0.220	0.415
	Went around the gate	0.140	0.350
	Went through the gate	0.060	0.240
	Other actions	0.140	0.336
Driver characteristics	Middle driver	0.430	0.495
	Old driver	0.320	0.467
	Young driver	0.250	0.434
	Female driver	0.260	0.439
	Male driver	0.740	0.439
	Driver was not in vehicle	0.160	0.364
	Driver was in vehicle	0.840	0.364
Vehicle characteristics	Auto	0.530	0.500
	Bus	0.000	0.036
	Motorcycle	0.010	0.077
	Truck	0.350	0.477
	Van	0.020	0.144
	Other vehicles	0.100	0.294
Environmental characteristics	Day	0.560	0.496
	Dusk	0.110	0.314
	Dark	0.240	0.427
	Dawn	0.090	0.280
	Clear	0.710	0.456
	Cloudy	0.210	0.406
	Fog	0.010	0.093
	Rain	0.060	0.237
	Sleet	0.000	0.045
	Snow	0.020	0.125

* Adverse crossing surface refers to wet, icy, snowy, slushy, sandy, muddy, or greasy roadway surfaces. Obstructed view is caused by highway vehicles, passing train, standing railroad equipment, permanent structure, topography, vegetation, or other entities.

4. Method

The random parameters logit model with heterogeneity in the means and variances is utilized in this study to investigate the non-divided two-way HRGC injury severity. By allowing relevant parameters to vary across individuals, the underlying mechanism and the unobserved heterogeneity that drives traffic accident outcomes can be explicitly uncovered. To determine the specific injury severity level n (n = 1-Uninjured, 2-Injured, or 3-Killed) of observation k, the linear utility function is defined as

$$U_{kn} = \boldsymbol{\beta}_k \boldsymbol{X}_{kn} + \varepsilon_{kn} \tag{1}$$

where U_{kn} is the severity function that determines the probability of injury severity n in crash k, $\boldsymbol{\beta}_k$ is the vector of parameters estimated for injury severity level n in observation k which is allowed to vary randomly across observations in the mixed logit models, $\boldsymbol{X}_{kn}$ is the vector of explanatory variables influencing HRGC injury severities, and ε_{kn} is a stochastic error term.

If the disturbance term is assumed to be generalized extreme value distributed, which allows parameters to vary across observations, a standard multinomial logit model can be defined as (McFadden, 1981)

$$P_{kn} = \int \frac{\exp\left(\boldsymbol{\beta}_k \boldsymbol{X}_{kn}\right)}{\sum \exp\left(\boldsymbol{\beta}_k \boldsymbol{X}_{kn}\right)} f(\boldsymbol{\beta}|\boldsymbol{\varphi}) d\boldsymbol{\beta} \tag{2}$$

where P_{kn} is the probabilities of injury severity outcome n in crash k, $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$ is the probability density function of vtor $\boldsymbol{\beta}$ and $\boldsymbol{\varphi}$ is a vector of parameters describing the density function, while all other terms are defined previously.

When the heterogeneity in means and variances exists among random parameters, the vector $\boldsymbol{\beta}_{kn}$ that allows estimated parameters to vary across accidents can be defined as (Ahmed et al., 2023; Behnood and Mannering, 2017)

$$\boldsymbol{\beta}_{kn} = \boldsymbol{\beta}_k + \boldsymbol{\delta}_{kn} \boldsymbol{Z}_{kn} + \sigma_{kn} \exp\left(\boldsymbol{\omega}_{kn} \boldsymbol{W}_{kn}\right) \upsilon_{kn} \tag{3}$$

where $\boldsymbol{\beta}_k$ represents the mean parameters estimated across all crashes, $\boldsymbol{Z}_{kn}$ is the vector of crash-specific variables with the heterogeneity in the mean that determines drivers' injury severity n, $\boldsymbol{\delta}_{kn}$ is a corresponding vector of estimable parameters, $\boldsymbol{W}_{kn}$ is another vector of crash-specific explanatory variables explaining the heterogeneity in the standard deviation σ_{kn} with corresponding parameter vector $\boldsymbol{\omega}_{kn}$, and υ_{kn} is the error term.

To estimate the random parameters logit model and obtain accurate estimates of the parameters, a simulation-based maximum likelihood method with 1000 Halton draws is used in this study (Halton, 1960; Islam et al., 2020; Yan et al., 2022). Previous studies have shown that the normal distribution is the best-suited distribution to describe the central tendency and variations of random variables in crash injury severity modeling (Ahmed et al., 2023; Pervez et al., 2022; Waseem et al., 2019). Thus, the random parameters in this study are assumed to follow the normal distribution. To provide an explicit interpretation of the impacts of significant variables X_{kn} on the probabilities of injury outcomes n, marginal effects of the explanatory variables X_{kn} on n-th crash injury are calculated as well (Washington et al., 2020)

$$\frac{\partial P_n(k)}{\partial X_{kn}} = \overline{P}_n(k) \left[\text{given } X_{kn} = 1 \right] - \overline{P}_n(k) \left[\text{given } X_{kn} = 0 \right] \tag{4}$$

To evaluate the model' performance, four different criteria including log-likelihood, McFadden Pseudo R-squared ρ^2 values, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) are used:

$$\rho^2 = 1 - LL(\beta)/LL(0) \tag{5}$$

$$AIC = -2LL(\beta) + 2K \tag{6}$$

$$BIC = -2LL(\beta) + K * \ln(N) \tag{7}$$

where $LL(\beta)$ is the log-likelihood at convergence, $LL(0)$ is the log-likelihood at zero, K is the number of parameters in the estimated model, and N is the number of observations.

Specifically, log-likelihood is a measure of how well the model describes the observed data, with higher values indicating a better fit. The ρ^2 provides a measure of the proportion of variance explained by the model, with a higher value indicating a better fit. The AIC and BIC are measures of model complexity, with a lower value indicating a better tradeoff between model fit and complexity (Akaike, 1974; Schwarz, 1978).

5. Results and discussions

The estimation results of the random parameters logit model with heterogeneity in means are presented in Table 2, which indicates an overall satisfactory statistical fit with ρ^2 values being 0.364. All the variables listed in Table 2 are identified as significant contributors to one category of crash severity at 99%, 95%, or 90% confidence level. This model also obtains suitable AIC and BIC values. The marginal effects of the significant explanatory factors (arranged by variable classification) affecting crash severities are shown in Fig. 2, where a positive value indicates that a one-unit increase in an explanatory variable increased the severity outcome probabilities, while a negative value decreases.

5.1. Heterogeneity in the means of random parameters

Table 2 shows that the did not stop variable specific to injured category is identified as the random parameter with normal distribution. The findings indicate that a varying influence of the unobserved characteristics on the injury severities is captured by this random parameter. The probability density diagram of the random parameter is presented in Fig. 3. The impact of all explanatory

Table 2
Model estimation results of random parameters logit model with heterogeneity in means.

Variables	Coefficient	t-stat	Marginal Effects		
			Uninjured [UI]	Injured [I]	Killed [K]
Constant [I]	−3.780***	−10.00			
Constant [K]	−4.949***	−11.79			
Random parameter (normally distributed)					
Did not stop [I]	1.351**	2.50	−0.0283	0.0307	−0.0024
Standard deviation of parameter distribution	3.195**	2.32			
Heterogeneity in the means of random parameter					
Did not stop: Dark [I]	−1.313*	−1.90			
Did not stop: Land with commercial power [I]	−1.867**	−2.32			
Crossing characteristics					
Highway speed limit > 25 MPH [K]	−0.420*	−1.84	0.0122	0.0049	−0.0172
Industry track [UI]	1.143**	2.54	0.0063	−0.0051	−0.0013
Yard track [UI]	1.233***	3.05	0.0072	−0.0059	−0.0013
Adverse crossing surface [UI]	0.473**	2.09	0.0085	−0.0064	−0.0021
Crash characteristics					
Estimated vehicle speed > 25 MPH [UI]	−0.912***	−3.43	−0.0128	0.0082	0.0046
Train speed > 45 MPH [K]	1.728***	7.84	−0.0282	−0.0125	0.0407
Went around the gate [UI]	−0.607***	−3.54	−0.0198	0.0148	0.0050
Driver characteristics					
Old driver [UI]	−0.443***	−3.00	−0.0239	0.0170	0.0069
Female driver [I]	0.443**	2.46	−0.0142	0.0165	−0.0023
Driver was in vehicle [UI]	−2.782***	−7.43	−0.4092	0.2943	0.1149
Vehicle characteristics					
Motorcycle [K]	2.110**	2.49	−0.0014	−0.0008	0.0022
Truck [UI]	0.646***	3.98	0.0297	−0.0208	−0.0089
Model statistics					
Number of observations (N)	1503				
Number of estimated parameters (K)	18				
Log-likelihood at zero, $LL(0)$	−1651.214				
Log-likelihood at convergence, $LL(\beta)$	−1050.897				
AIC	2137.8				
BIC	2233.5				
$\rho^2 = 1 - LL(\beta)/LL(0)$	0.364				

***,**,*represents significance at 99%, 95%, 90% level.

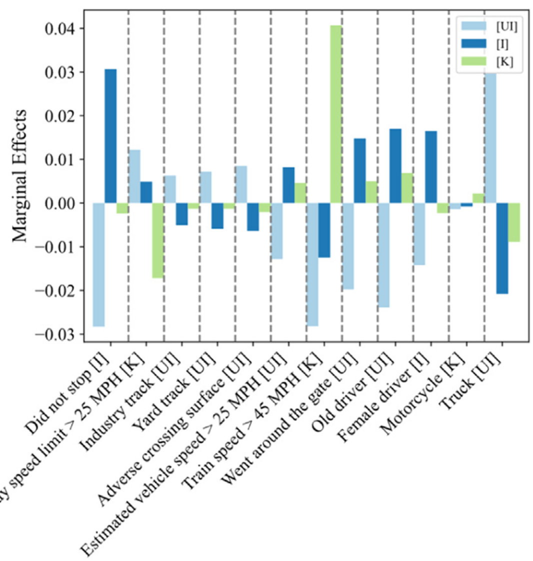

Fig. 2. Marginal effects of explanatory variables.

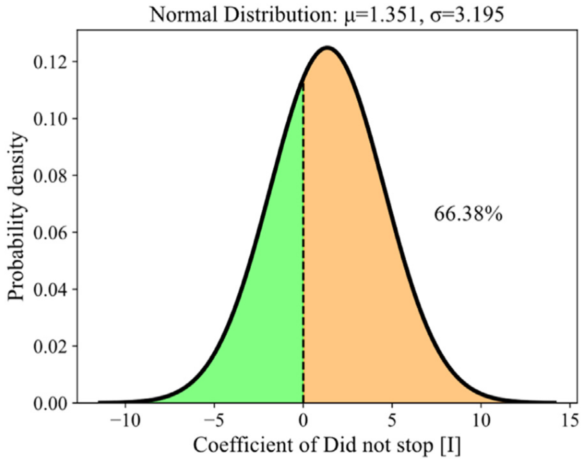

Fig. 3. Distribution of the random parameter.

variables on the means and variances of the random parameters is thoroughly assessed to determine the presence of heterogeneity, and it is found that only the heterogeneity in the means is significant. The present study reveals that the did not stop indicator defined to injured category produces a random parameter that is normally distributed with a mean of −3.780 and a standard deviation of −4.949. This implies that this variable decreases the likelihood of injured outcomes for 33.62% of the HRGC crashes while increasing the likelihood of that for 66.38% of these observations.

Regarding the heterogeneity in means, the dark and the land with commercial power indicators are observed to decrease the means of heterogeneity in the did not stop indicator for injured outcomes, suggesting there may be an increase in severe injuries. In the context of HRGC crashes, this result is intuitive since the unique driving condition of the crossings may substantially affect the drivers' driving performance under the dark and the land with commercial power indicator (Diaz-Piedra et al., 2021; Dong et al., 2011; Fountas et al., 2020). The inadequate lighting conditions and mixed traffic flow in the commercial district could bring serious obstacles to drivers in dealing with complicated traffic driving scenarios (e.g., the high train speed, obstructed view, or improper post-accident treatment) and consequently lead to more serious injury severities in such crashes.

5.2. Crossing characteristics

The highway speed limit > 25 MPH variable is a significant factor affecting injury severity, with a negative coefficient of −0.420 in the killed model. The marginal effects imply that this variable may be less likely to result in fatalities (by 0.0459), while it is associated with a higher probability of uninjured and injured outcomes (by 0.0122 and 0.0049, respectively). This could be because that higher speed limits on highways are associated with better-designed crossings that have better warning systems (Chen et al., 2022). Another explanation could be that drivers are more likely to be aware of the potential risks with the presence of speed limits, which decreases the likelihood of fatal injuries (Fountas et al., 2021).

The indicators for industry tracks and the yard tracks are both positively correlated with being uninjured in a railway crossing crash but negatively correlated with being injured and killed, with marginal effects of 0.0063 and 0.0072 for the uninjured category, −0.0051 and −0.0059 for the injured category, while −0.0013 and −0.0013 for the killed category, respectively. One possible reason is that industry tracks and yard tracks typically have lower traffic volumes compared to mainline tracks, which could result in a reduced likelihood of collisions between trains and vehicles (Lin et al., 2023). What's more, the maintenance of crossing infrastructures may be more frequent on industry and yard tracks, thus reducing the risk of accidents caused by malfunctioning crossing signals or barriers (Zhang et al., 2022).

The coefficient estimates for the variable representing the adverse crossing surface is positive, indicating that such road surface may increase the likelihood of being uninjured in a railway crossing crash (by 0.0085), but decrease the likelihood of being injured and killed (by 0.0064 and 0.0021, respectively). It is noteworthy that adverse crossing surfaces may result in reduced driving speeds and increased caution among drivers, potentially lowering the likelihood of severe accidents (Pang et al., 2022; Yan et al., 2022). Nevertheless, it is important to note that contradictory results have also been found in previous studies and the discrepancy may be attributed to the higher frequency of driving errors and dangerous driving behaviors in unfavorable road conditions (Alnawmasi and Mannering, 2022; Wang et al., 2022). As a matter of fact, the relationship between road surface conditions and injury severity can be complex and can vary depending on several factors, such as weather conditions, traffic volume, and driver behavior.

5.3. Crash characteristics

The estimated vehicle speed greater than 25 MPH indicator is found to be negatively correlated with being uninjured in a railway crossing crash, with a marginal effect of −0.0128, while being positively correlated with being injured and being killed with a marginal effect of 0.0082 and 0.0046. Meanwhile, the train speed greater than 45 MPH indicator is also found to be positively correlated with the probability of being killed in a railway crossing crash, with a marginal effect of 0.0407, while being negatively correlated with being uninjured with a marginal effect of −0.0282 and being injured with a marginal effect of −0.0125. This could be because higher speeds lead to less time for the vehicle to react, thus reducing the probability of avoiding a collision (Ahmed et al., 2023; Hao and Daniel, 2014). Additionally, higher train speeds could result in more severe impacts that are more likely to cause fatal injuries to vehicle occupants (Hao and Daniel, 2013). It is important to note that this result emphasizes the importance of driving at reduced speed at railway crossings to avoid potential collisions with trains.

Moreover, the indicator for vehicles that went around the gate is found to be negatively correlated with being uninjured in a railway crossing crash, with a marginal effect of −0.0198, while being positively correlated with being injured and being killed with a marginal effect of 0.0148 and 0.0050, respectively. In fact, vehicles that go around the gate indicate a disregard for safety protocols, suggesting that the driver may be engaging in other risky behaviors and consequently increase the likelihood of being injured or killed in a crash (Ahmed et al., 2023).

5.4. Driver characteristics

It is found that being an old driver increases the probability of being injured by 0.0170 and the probability of being killed by 0.0069 while decreasing the probability of being uninjured by 0.0239. One possible explanation is that physical and cognitive decline associated with aging may hinder an older driver's capacity to respond effectively to unexpected events at the HRGCs (Ahmed et al., 2023; Alnawmasi and Mannering, 2022). Furthermore, older drivers may have a higher prevalence of preexisting health conditions that can exacerbate injuries sustained in the crash, contributing to an elevated probability of being injured or killed (Lynch et al., 2022).

The female driver indicator has a positive coefficient of 0.443 for the injured category. The marginal effects further indicate that female drivers are more likely to be involved in more serious injury crashes (by 0.0142) due to their susceptibility to traffic accidents (Islam and Mannering, 2020). This finding may be attributed to the difference in driving behavior and patterns between male and female drivers (Yan et al., 2022). Additionally, the lack of experience, coupled with less exposure to dangerous driving situations, may also lead to a higher likelihood of mishandling these situations and ultimately result in more severe injuries (Tement et al., 2020).

The indicator variable for drivers who were in the vehicle during the crash is found to decrease the probability of being uninjured by 0.4092 while increasing the probability of being injured and killed by 0.2943 and 0.1149, respectively. It may be due to the fact that drivers who are inside the vehicle during a crash have a higher likelihood of being directly impacted by the collision, and therefore are more likely to sustain injuries or fatalities compared to those who are not in the vehicle (Ahmed et al., 2023; Mathew and Benekohal, 2021). This indicator may also reflect situations where the driver is unable to react in time to prevent or minimize the impact of the collision, leading to a higher probability of being injured or killed.

5.5. Vehicle characteristics

The type of vehicles can have substantial impacts on the severity of injuries sustained by the individual involved at non-divided two-way traffic HRGCs. Specifically, the motorcycle involvement indicator is found to be positively correlated with killed outcomes, with the marginal effects suggesting that this variable increases the probability of being killed in a railway crossing accident by 0.0022 while reducing the probability of being injured and uninjured by 0.0008 and 0.0014, respectively. This could be due to the vulnerability of motorcyclists in accidents, as they lack the protective structure of a vehicle and are more likely to sustain fatal injuries in the HRGC collision (Ahmed et al., 2023; Das, 2021). Meanwhile, the smaller size and higher speed of motorcycles relative to other vehicles could make them harder to be detected and more prone to accidents at HRGCs (Li et al., 2021). The result also indicates that truck involvement is positively associated with the uninjured category. It is found that this indicator increases the probability of being uninjured by 0.0297 while decreasing the probability of being injured and killed by 0.0208 and 0.0089, respectively. One possible reason could be that trucks are often operated by skilled drivers who adhere to safety regulations and exercise caution when crossing railway tracks (Doubek et al., 2021). Another possible explanation is that trucks are larger and more durable vehicles that may provide better protection to their occupants (Tyndall, 2021), potentially reducing the risk of sustaining injuries in HRGC crashes.

6. Conclusions

Highway-rail grade crossing crashes in non-divided two-way traffic scenarios are a significant public concern and have caused numerous fatalities and injuries over the years. Using the HRGC crash dataset collected in the United States from 2019 to 2020, this study aims to identify the important factors and their heterogeneous impacts that make the injury severity in non-divided two-way traffic HRGC crashes more severe. The injury outcomes were divided into three categories: Uninjured, Injured, and Killed. A random parameters logit model with heterogeneity in means was employed in this paper considering crossings, crashes, drivers, vehicles, and environmental characteristics. Additionally, the marginal effects were also investigated to gain a better insight into the effects of these contributing factors on the severity of HRGC crashes.

The present study found that the did not stop behavior produced varying effects on injury severity outcomes. Therefore, there is a great necessity for targeted programs to educate drivers about avoiding risky or improper driving behaviors at HRGCs. Besides, clear signage, visible traffic signals, and proper lane markings should be incorporated into the crossing to enhance visibility and provide clear instructions, thus reducing the likelihood of non-stop behavior. In this study, the random parameters' means were observed to be associated with the dark and land with commercial power indicators. Installing adequate lights in roadways and crossings is necessary to improve visibility and reduce the risk of accidents. And encouraging the use of reflective materials on vehicles during nighttime travel can also be helpful. To enhance HRGC safety in commercial areas, it is crucial to promote the installation of warning signs that provide drivers with important information and alerts, including speed limits, pedestrian crossings, and potential hazards. This study further identified several factors that increase the likelihood of severe injuries in HRGC crashes, including estimated vehicle speed > 25 MPH, train speed > 45 MPH, going around the gate, old driver, female driver, motorcycle, and the driver was in vehicle indicators. Based on the results, countermeasures such as implementing lower speed limits at crossings, strengthening gate enforcement, and developing targeted initiatives for vulnerable road users could be taken to reduce the likelihood of injuries. Notably, the adverse crossing surface and truck indicators exhibited unexpected marginal effects and decreased the possibility of severe injury outcomes at non-divided two-way traffic HRGCs. Nevertheless, it is essential to enhance the maintenance of HRGCs, including repairing potholes, improving road surfaces, and ensuring proper drainage to minimize the risk under adverse crossing surface conditions. Additionally, educational programs should be implemented to raise awareness among truck drivers about the importance of following traffic rules, yielding to trains, and exercising caution at crossings.

It should be noted that our findings and conclusions are based on the variables provided in the historical crash records, and the absence of some variables such as traffic parameters (e.g., real-time traffic speed, acceleration/deceleration) and location-specific context (e.g., crossing-intersection distance, crossing angle, number of main tracks), could potentially limit the generalizability of the result. Besides, under-reporting of non-injury crashes could bias the results of parameter estimation (Mannering, 2018). In the future, databases with more detailed information on HRGC crashes should be constructed to further explore the unobserved heterogeneity (Ahmed et al., 2023; Mannering, 2018). Besides, more advanced methodologies combining resampling methods should also be developed to accurately examine the heterogeneity in HRGC crashes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The work described in this paper was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 15221821) and the Research Committee of The Hong Kong Polytechnic University (UAKQ).

References

Ahmed, S.S., Corman, F., Anastasopoulos, P.Ch., 2023. Accounting for unobserved heterogeneity and spatial instability in the analysis of crash injury-severity at highway-rail grade crossings: a random parameters with heterogeneity in the means and variances approach. Anal. Method. Accid. Res. 37, 100250. doi:10.1016/j.amar.2022.100250.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 196, 716–723. doi:10.1109/TAC.1974.1100705.

Alnawmasi, N., Mannering, F., 2022. A temporal assessment of distracted driving injury severities using alternate unobserved-heterogeneity modeling approaches. Anal. Method. Accid. Res. 34, 100216. doi:10.1016/j.amar.2022.100216.

Alnawmasi, N., Mannering, F., 2019. A statistical assessment of temporal instability in the factors determining motorcyclist injury severities. Anal. Method. Accid. Res. 22, 100090. doi:10.1016/j.amar.2019.100090.

Behnood, A., Mannering, F., 2017. Determinants of bicyclist injury severities in bicycle-vehicle crashes: a random parameters approach with heterogeneity in means and variances. Anal. Method. Accid. Res. 16, 35–47. doi:10.1016/j.amar.2017.08.001.

Calvi, A., Cafiso, S.D., D'Agostino, C., Kieć, M., Petrucci, G., 2023. A driving simulator study to evaluate the effects of different types of median separation on driving behavior on 2 + 1 roads. Accid. Anal. Prevent. 180, 106922. doi:10.1016/j.aap.2022.106922.

Chen, S., Chen, Y., Xing, Y., 2022. Comparison and analysis of crash frequency and rate in cross-river tunnels using random-parameter models. J. Transp. Saf. Secur. 142, 280–304. doi:10.1080/19439962.2020.1779420.

Das, S., 2021. Identifying key patterns in motorcycle crashes: findings from taxicab correspondence analysis. Transportmetrica A: Transp. Sci. 174, 593–614.

Diaz-Piedra, C., Rieiro, H., Di Stasi, L.L., 2021. Monitoring army drivers' workload during off-road missions: an experimental controlled field study. Saf. Sci. 134, 105092. doi:10.1016/j.ssci.2020.105092.

Dong, Y., Hu, Z., Uchimura, K., Murayama, N., 2011. Driver inattention monitoring system for intelligent vehicles: a review. IEEE Trans. Intell. Transp. Syst. 12 (2), 596–614. doi:10.1109/TITS.2010.2092770.

Doubek, F., Salzmann, F., de Winter, J., 2021. What makes a good driver on public roads and race tracks? An interview study. Transp. Res. Part F: Traffic Psychol. Behav. 80, 399–423. doi:10.1016/j.trf.2021.04.019.

Federal Railroad Administration, 2021. Highway/Rail grade crossing incidents. URL https://railroads.dot.gov/accident-and-incident-reporting/highwayrail-grade-crossing-incidents/highwayrail-grade-crossing (accessed 3.1.23).

Fountas, G., Fonzone, A., Gharavi, N., Rye, T., 2020. The joint effect of weather and lighting conditions on injury severities of single-vehicle accidents. Anal. Method. Accid. Res. 27, 100124. doi:10.1016/j.amar.2020.100124.

Fountas, G., Fonzone, A., Olowosegun, A., McTigue, C., 2021. Addressing unobserved heterogeneity in the analysis of bicycle crash injuries in Scotland: a correlated random parameters ordered probit approach with heterogeneity in means. Anal. Method. Accid. Res. 32, 100181. doi:10.1016/j.amar.2021.100181.

Gao, L., Lu, P., Ren, Y., 2021. A deep learning approach for imbalanced crash data in predicting highway-rail grade crossings accidents. Reliab. Eng. Syst. Saf. 216, 108019. doi:10.1016/j.ress.2021.108019.

Ghomi, H., Bagheri, M., Fu, L., Miranda-Moreno, L.F., 2016. Analyzing injury severity factors at highway railway grade crossing accidents involving vulnerable road users: a comparative study. Traffic Inj. Prev. 17 (8), 833–841. doi:10.1080/15389588.2016.1151011.

Halton, J.H., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. 2, 84–90.

Hao, W., Daniel, J., 2014. Motor vehicle driver injury severity study under various traffic control at highway-rail grade crossings in the United States. J. Saf. Res. 51, 41–48. doi:10.1016/j.jsr.2014.08.002.

Hao, W., Daniel, J.R., 2013. Severity of injuries to motor vehicle drivers at highway–rail grade crossings in the United States. Transp. Res. Rec. 23841, 102–108. doi:10.3141/2384-12.

Hao, W., Kamga, C., Daniel, J., 2015. The effect of age and gender on motor vehicle driver injury severity at highway-rail grade crossings in the United States. J. Saf. Res. 55, 105–113. doi:10.1016/j.jsr.2015.08.006.

Islam, M., Alnawmasi, N., Mannering, F., 2020. Unobserved heterogeneity and temporal instability in the analysis of work-zone crash-injury severities. Anal. Method. Accid. Res. 28, 100130.

Islam, M., Mannering, F., 2020. A temporal analysis of driver-injury severities in crashes involving aggressive and non-aggressive driving. Anal. Method. Accid. Res. 27, 100128. doi:10.1016/j.amar.2020.100128.

Jamal, A., Zahid Khattak, M., Rahman, M.T., Hasan, A., Almoshaogeh, M., Farooq, D., Ahmad, M., 2021. Injury severity prediction of traffic crashes with ensemble machine learning techniques: a comparative study. Int. J. Injury Control Saf. Promot. doi:10.1080/17457300.2021.1928233.

Keramati, A., Lu, P., Iranitalab, A., Pan, D., Huang, Y., 2020. A crash severity analysis at highway-rail grade crossings: the random survival forest method. Accid. Anal. Prevent. 144, 105683. doi:10.1016/j.aap.2020.105683.

Kutela, B., Kidando, E., Kitali, A.E., Mwende, S., Langa, N., Novat, N., 2022. Exploring pre-crash gate violations behaviors of drivers at highway-rail grade crossings using a mixed multinomial logit model. Int. J. Injur. Control Saf. Promot. 29 (2), 226–238. doi:10.1080/17457300.2021.1990348.

Li, J., Fang, S., Guo, J., Fu, T., Qiu, M., 2021. A Motorcyclist-injury severity analysis: a comparison of single-, two-, and multi-vehicle crashes using latent class ordered probit model. Accid. Anal. Preven. 151, 105953. doi:10.1016/j.aap.2020.105953.

Lin, C.-Y., Blumenfeld, M., Gerstein, T., Barkan, C.P.L., Jack, A., Abdurrahman, U.T., 2023. International benchmarking of railroad safety data systems and performance – a cross-continental case study. J. Rail Transp. Plann. Manag. 26, 100384. doi:10.1016/j.jrtpm.2023.100384.

Liu, P., Fan, W., 2021. Analysis of head-on crash injury severity using a partial proportional odds model. J. Transp. Saf. Secur. 137, 714–734. doi:10.1080/19439962.2019.1667933.

Liu, R., Yan, X., Ma, S., Xue, Q., 2022. Eye movement as a function to explore the effects of improved signs design and audio warning on drivers' behavior at STOP-sign-controlled grade crossings. Accid. Anal. Prevent. 172, 106693. doi:10.1016/j.aap.2022.106693.

Lombardi, D.A., Horrey, W.J., Courtney, T.K., 2017. Age-related differences in fatal intersection crashes in the United States. Accid. Anal. Prevent. 99, 20–29. doi:10.1016/j.aap.2016.10.030.

Lynch, S.D., Weaver, A.A., Barnard, R.T., Kiani, B., Stitzel, J.D., Zonfrillo, M.R., 2022. Age-based differences in the disability of spine injuries in pediatric and adult motor vehicle crash occupants. Traffic Inj. Prev. 236, 358–363. doi:10.1080/15389588.2022.2086980.

Mahmud, S.M.S., Ferreira, L., Hoque, Md.S., Tavassoli, A., 2022. Overtaking risk modeling in two-lane two-way highway with heterogeneous traffic environment of a low-income country using naturalistic driving dataset. J. Saf. Res. 80, 380–390. doi:10.1016/j.jsr.2021.12.019.

Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Anal. Method. Accid. Res. 17, 1–13. doi:10.1016/j.amar.2017.10.002.

Mathew, J., Benekohal, R.F., 2021. Highway-rail grade crossings accident prediction using Zero inflated negative binomial and empirical bayes method. J. Saf. Res. 79, 211–236. doi:10.1016/j.jsr.2021.09.003.

McFadden, D., 1981. Econometric models of probabilistic choice. Struct. Anal. Discr. Econom. Appl., 198272.

Pang, J., Krathaus, A., Benedyk, I., Ahmed, S.S., Anastasopoulos, P.Ch., 2022. A temporal instability analysis of environmental factors affecting accident occurrences during snow events: the random parameters hazard-based duration model with means and variances heterogeneity. Anal. Method. Accid. Res. 34, 100215. doi:10.1016/j.amar.2022.100215.

Pervez, A., Lee, J., Huang, H., 2022. Exploring factors affecting the injury severity of freeway tunnel crashes: a random parameters approach with heterogeneity in means and variances. Accid.Anal. Prevent. 178, 106835. doi:10.1016/j.aap.2022.106835.

Schwarz, G., 1978. Estimating the Dimension of a Model. The Annal. Stat. 62, 461–464. doi:10.1214/aos/1176344136.

Soleimani, S., Mousa, S.R., Codjoe, J., Leitner, M., 2019. A comprehensive railroad-highway grade crossing consolidation model: a machine learning approach. Accid. Analy. Prevent. 128, 65–77. doi:10.1016/j.aap.2019.04.002.

Tement, S., Plohl, N., Horvat, M., Musil, B., Jakus, G., Sodnik, J., 2020. Driving demands, stress reactivity and driving behavior: an interactional approach. Transp. Res. Part F: Traffic Psychol. Behav. 69, 80–90. doi:10.1016/j.trf.2020.01.001.

Tjahjono, T., Kusuma, A., Pratiwi, Y.Y., Purnomo, R.Y., 2019. Identification determinant variables of the injury severity crashes at road-railway level crossing in indonesia. In: Transportation Research Procedia, 21st EURO Working Group on Transportation Meeting, 37. Braunschweig, Germany, pp. 211–218. doi:10.1016/j.trpro.2018.12.185 EWGT 2018, 17th –19th September 2018.

Tyndall, J., 2021. Pedestrian deaths and large vehicles. Econom. Transp. 26–27, 100219. doi:10.1016/j.ecotra.2021.100219.

Wang, C., Chen, F., Zhang, Y., Wang, S., Yu, B., Cheng, J., 2022. Temporal stability of factors affecting injury severity in rear-end and non-rear-end crashes: a random parameter approach with heterogeneity in means and variances. Anal. Method. Accid. Res. 35, 100219. doi:10.1016/j.amar.2022.100219.

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. Accid. Anal. Prevention 123, 12–19. doi:10.1016/j.aap.2018.10.022.

Washington, S., Karlaftis, M., Mannering, F., Anastasopoulos, P., 2020. Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC.

Wu, L., Shen, Q., Li, G., 2022. Identifying risk factors for autos and trucks on highway-railroad grade crossings based on mixed logit model. Int. J. Environ. Res. Public Health 19 (22), 15075. doi:10.3390/ijerph192215075.

Yan, D., Li, K., Zhu, Q., Liu, Y., 2023. A railway accident prevention method based on reinforcement learning – active preventive strategy by multi-modal data. Reliab. Eng. Syst. Saf. 234, 109136. doi:10.1016/j.ress.2023.109136.

Yan, X., He, J., Wu, G., Zhang, C., Wang, C., Ye, Y., 2022. Differences of overturned and hit-fixed-object crashes on rural roads accompanied by speeding driving: accommodating potential temporal shifts. Anal. Method. Accid. Res. 35, 100220. doi:10.1016/j.amar.2022.100220.

Zhang, Z., Zaman, A., Xu, J., Liu, X., 2022. Artificial intelligence-aided railroad trespassing detection and data analytics: methodology and a case study. Accid. Anal. Prevent. 168, 106594. doi:10.1016/j.aap.2022.106594.