

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

Hamid Shamanian Esfahani^{1,*}, Seyed Teimour Hoseini², Reza Sakhaei², Majid Ehsani Sohi²

Received: 2024/04/26

Accepted: 2024/08/26

Abstract

The high rate of traffic crashes and consequent losses have caused many social and economic problems in developing countries like Iran. According to previous studies, 60-70% of rural damages in Iran occur in marginal areas around cities. Hence, identifying effective factors in rural crashes around cities is of great importance to take proper safety measures in such areas. This requires developing exclusive models to predict rural crash frequency around cities. Given the limitations of fixed-parameters models, such as Poisson and negative binomial regression, the present study uses a random-parameter approach, including the random parameters negative binomial (RPNB) and random effects negative binomial (RENB) models to determine the factors affecting rural crash frequency at marginal areas around cities. The data on rural crashes on roads leading to Isfahan over 3 years, were used to develop the models. The random-parameter models were used to consider the unobserved heterogeneity among crash observations. The goodness-of-fit measures indicated the acceptable accuracy of the developed models. Meanwhile, the likelihood ratio test revealed a slight difference between the RPNB and RENB. According to the outcomes, the factors influencing crashes around cities included traffic volume, the ratio of local and non-local vehicles, the proportion of longitudinal distance-limit violations, access points controlling, and access to rest area facilities.

Keywords: Marginal areas around cities, crash frequency modelling, unobserved heterogeneity, random parameters negative binomial (RPNB) model, random effects negative binomial (RENB) model

* Corresponding author. E-mail: hamid.shamanian@srbiau.ac.ir

¹ Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

² Amin University, Tehran, Iran

1. Introduction

The mortality rate of traffic crashes in Iran is 5.20 per 100,000 population (WHO, 2018). According to the statistics published by the Road Maintenance and Transportation Organization of Iran (RMTO, 2022), over the past ten years, 65% of the losses caused by crashes were due to rural ones. On the other hand, previous studies on the spatial analysis of rural crashes in the country indicate that a considerable portion of rural crashes have occurred in marginal areas around cities (Sajed, Shafabakhsh, and Bagheri, 2019; Elyasi, Saffarzadeh, and Boroujerdian, 2018; Shafabakhsh, Famili, and Akbari, 2016; Elyasi, Saffarzadeh, and Boroujerdian, 2016; Mohaymany, Shahri, and Mirbagheri, 2013; Effati, Rajabi, and Samadzadegan, 2012; Boroujerdian, Saffarzadeh, and Abolhasannejad, 2010). Based on the RMTO crash data, 67% of rural crashes have been related to marginal areas. Considering the average number of rural crashes in the past ten years equal to 122,100 (RMTO, 2022), the number of crashes at marginal areas around cities in Iran equaled 81,800 on average in these years. In a study on the effect of marginal areas around cities on rural crashes, Shamanian Esfahani, Afandizadeh, and Naderan (2022) introduced the boundary of influence area (BIA) around cities. According to the study, BIA (which determines the marginal road section in this study) for a road leading to a city indicates a boundary on the road at a particular distance from the city where rural crash frequency changes significantly. The difference in the rural crash frequency in marginal and basic highway segments can be due to the different characteristics of highway features (such as the distance and density of access points) or different environmental road characteristics (for example, the length of adjacent land use or type of land use) of such areas compared to the basic highway segments. This can also be caused by the different

behaviors of drivers in marginal areas around cities and basic highway segments. Given the high portion of crashes around cities from all rural crashes, determining the major parameters of crashes in these areas by developing exclusive crash frequency prediction models is of great importance.

The Poisson regression and negative binomial (NB) models are the traditional models used for estimating the crash frequency. These models assume a fixed relationship between crash frequency and factors effective in crashes in temporal and spatial dimensions. However, this assumption is mostly rejected since crash observations depend on time and/or space (Han et al., 2018). On the other hand, the number of factors affecting the occurrence of crashes, including the parameters describing the behavior of humans, characteristics of the vehicles and roads, traffic characteristics, and environmental conditions, is very high. Meanwhile, crash databases, which are typically extracted from reports provided by the traffic police make up a small portion of the factors effective in the occurrence of crashes. Accordingly, many factors are neglected in the modeling process. For instance, the weather conditions are variable over time (for several years); thus, the drivers' reactions to these conditions would be variable. As another example, the monitoring level of the police varies for different roads, which can impact the probability of crashes on different roads. If the unobserved variables have high correlation with the variables used in the model, ignoring them can lower the accuracy of the developed models and result in errors in the estimation of the crashes made by them (Mannering et al., 2020; Mannering, Shankar, and Bhat, 2016; Mannering and Bhat, 2014). This intrinsic defect of crash databases is called unobserved heterogeneity.

The crash modeling methods aimed at eliminating the temporal and spatial dependence of crash observations and errors caused by the unobserved heterogeneity have

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

been improved. The random-parameter approach is one of such improved methods in modeling. To take into account the heterogeneity between observations of different classes (e.g., crashes in the segments of two different roads or the intersections of two different cities), this approach allows the parameters estimated in the model to be different for observations of various classes. This procedure requires determining a continuous distribution for the desired descriptive variable by the analyzer.

A review of the literature demonstrates that most of the studies conducted in Iran have used traditional methods to create models for the estimation of the crash frequency while ignoring the errors caused by temporal and spatial dependences and unobserved heterogeneity (Faezi and Elyasi, 2020; Imaninasab, Sekhavati, and Hajihoseinloo, 2016; Nassiri, Najafi, and Mohamadian Amiri, 2014). On the other hand, given the cities' marginal areas' considerable portion of the whole road crashes in Iran, the development of no separate models to estimate the crashes of these areas can be considered another drawback of the previous studies. According to the explanations provided, the present study seeks to develop improved models for the prediction of crash frequency in the marginal areas of cities. To this end, the RPNB and RENB models were employed.

In the following section, previous studies modeling crash frequency are investigated. In the next step, the methodology of the present study, as well as the RPNB and RENB models, are explained. Then, some explanations are provided about the database used in the study and how the data were collected. At the end, the modeling results are described to conclude.

2. Literature Review

This section reviews the factors effective in the occurrence of crashes and the crash frequency analysis methods in previous studies.

2.1. Factors Effective in Crash Frequency

A crash frequency analysis requires identifying the major parameters affecting the occurrence of crashes and their effects on the crash frequency. The factors influencing crash frequency are classified into three main groups (Lord and Mannering, 2010): highway geometrical/traffic characteristics, behavioral parameters of drivers, and the spatial and temporal characteristics of the environment surrounding a road.

Afghari et al. (2018) introduced the number of lanes, lane width, shoulder width, highway median, and horizontal and vertical alignment of the road as the geometrical characteristics of highways. Many studies confirm the effects of these parameters on crash frequency (Wen et al., 2021; Rusli et al., 2017; Zou and Tarko, 2016; Dong et al., 2015; Ahmed et al., 2011; Fu et al., 2011; Zhang and Ivan, 2005; Lee and Mannering, 2002).

A large number of research articles have evaluated the impact of road traffic volume on crash frequency, with almost all of them emphasizing a direct relationship between the two parameters (Mousavi et al., 2019; Yu, Quddus, and Li, 2019; Ma et al., 2017; Anastasopoulos and Mannering, 2009; Donnel and Mason, 2006; Abdel-Aty and Radwan, 2000). There is also evidence proving that local drivers drive with less caution compared to non-local ones. Hence, the probability of crashes grows with the increased portion of local vehicles among the vehicles passing through a highway (Yu et al., 2019).

The unobserved heterogeneity in crash frequency analyses is mostly caused by the behavioral factors of drivers. Given the variability of these factors in different individuals, it is difficult to measure and record them in crash reports and introduce them in the crash frequency estimation models (Mannering, Shankar, and Bhat, 2016).

The road pavement and climate conditions are among the most important spatial factors whose

relationship with crash frequency has been investigated in some road safety studies. Most of these studies have reported the improper conditions of the road pavement and the climate as the reasons behind the increased crash frequency (Khoda Bakhshi and Ahmed, 2021; Park, Abdel-Aty, and Wang, 2017; Naik et al., 2016; Buddhavarapu, Banerjee, and prozzi, 2013; Abdel-Aty, Devarasetty, and Pande, 2009).

Some studies have assessed the relationship between the adjacent land use of the road and crash frequency. Wu, Song, and Meng (2021), Mohammadnazar, Arvin, and Khattak. (2021), and Kim, Brunner, and Yamashita. (2006) showed that commercial land use beside a highway increases rural crashes. This can be due to the changes in environmental conditions and a more complex relationship between road users and increased highway access points. These factors eventually increase the aggressive behaviors among drivers in these areas (Mohammadnazar, Arvin, and Khattak, 2021). Moreover, the evidence shows that the number of highway access points and the way they are controlled are related to the crash frequency. Investigating the factors affecting the occurrence of crashes in freeways in China, Hoe, Tarko, and Meng. (2018) showed that crashes are more likely in segments with interchanges than in basic segments. Mohammadnazar et al. (2021a) showed that the crash frequency decreases with the rise in the completely controlled access points. The full control of highway access points can reduce the convergence and divergence in the traffic flow, and thus, mitigate the adverse effects of access points on the main traffic flow (Mohammadnazar et al., 2021a). Investigating the factors effective in the occurrence of rural crashes in the entries of cities, Ehasni Sohi, Dashtestaninejad, and Khademi (2019) showed that crash frequency increases with the number of access points to the major roads in these areas.

2.2. Crash Frequency Analysis Approaches

Many of the studies developing crash frequency estimation models have employed traditional methods like Poisson regression (Ehasni Sohi, Dashtestaninejad, and Khademi, 2019; Hosseini, Mahdayi, and Nooghabi, 2018; Ayati and Abbasi, 2011; Ma, 2009; Kumara and Chin, 2005) and NB model (Faezi and Elyasi, 2020; Yan et al., 2020; Garach et al., 2016; Imaninasab, Sekhavati, and Hajihoseinloo, 2016; Nassiri, Najafi, and Mohamadian Amiri, 2014; Malyshkina and Mannering, 2010; Lord and Bonneson, 2007; Kim and Washington, 2006).

Neglecting to consider the potential unobserved heterogeneity (thus, implicitly assuming that the effects of observed factors are homogeneous across all observations) might generate biased and inconsistent parameter estimates, leading to incorrect conclusions and invalid crash estimations (Mannering, Shankar, and Bhat, 2016).

The random-parameters and random effects model formulations have been employed in conventional count models to account for the impacts of unobserved heterogeneity. Incorporating random effects into NB models aims to address unobserved temporal and spatial variations across different spatial locations and time periods (Gong et al., 2020). In the RENB model, it is assumed that the unobserved heterogeneity specific to individuals is entirely independent of the explanatory variable vector, with only the regression intercept being random (Moomen et al., 2020). In contrast to the RENB model, where only the intercept term is considered random, the RPNB model permits every estimated parameter to fluctuate among individual observations, encompassing unobserved heterogeneity across spatial and temporal dimensions (Hou, Tarko, and Meng, 2018).

Although some of the previous studies have compared random-parameter and fixed-

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

parameter models, proving the superiority of improved methods over the traditional ones (Shaon et al., 2018; Cai et al., 2018; Yu et al., 2019; Saeed et al., 2019; Park, Abdel-Aty, and Weng, 2016), the comparison between the RENB and RPNB models has received less attention. On the other hand, given the large portion of rural crashes in the marginal areas of cities in Iran, developing exclusive crash frequency models for these areas should be taken into account. Accordingly, the present study employs random-parameter approaches, the RENB and RPNB models in particular, to determine the factors affecting crash frequency in marginal areas around cities and compares the model results

3. Methodology

The random-parameter approach has been employed to incorporate unobserved heterogeneity into count models. The parameters of the RPNB model can be represented as:

$$\beta_{il} = b_l + \omega_{il} \quad (1)$$

The value of the estimated parameter for the explanatory variable l is represented by b_l , β_{il} denotes the parameter on explanatory variable l for the i th observation, and ω_{il} is an error term randomly distributed, assumed to follow a predefined distribution, such as normal, log-normal, uniform, or triangular. When the estimated standard deviation for a parameter significantly deviates from zero, it is treated as a random estimation; otherwise, it is regarded as a fixed variable.

Below are the equations providing the crash frequency parameter and the log-likelihood function:

$$\lambda_i | \omega_{il} = \exp(\beta_{il} x_i + \varepsilon_i) \quad (2)$$

$$LL = \sum_{i,l} \ln \int_{\omega_{il}} g(\omega_{il}) P(y_{il} | \omega_{il}) d\omega_{il} \quad (3)$$

In which, λ_i signifies the forecasted crash count in the i th observation, x_i refers to a vector containing explanatory variables, and $g(\omega_{il})$ represents the probability function of ω_{il} . Given the necessity for numerical integration with

random parameter distribution in both NB and Poisson models, the maximization of the log-likelihood function is achieved through a simulation-based maximum likelihood approach. Halton draws are utilized in the simulation approach, which has emerged as the predominant method for estimating random parameters and yields more precise results compared to conventional random drawings (Washington et al., 2020; Greene, 2012).

The Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and McFadden ρ^2 are applied to evaluate the goodness-of-fit of the model:

$$z = \beta X + \varepsilon \quad (4)$$

$$AIC = -2 LL(\beta) + 2k \quad (5)$$

$$BIC = -2LL(\beta) + k \ln(n) \quad (6)$$

in which $LL(\beta)$ represents the log-likelihood at convergence, $LL(0)$ denotes the log-likelihood with the constant term, k signifies the number of variables, and n stands for the number of observations. The optimal models are identified by their lowest BIC and AIC values, along with the highest log-likelihood value at convergence. The likelihood ratio test is employed to ascertain if the specification statistically surpasses the others:

$$\beta_{il} \quad \chi^2 \\ = b_l + \omega_{il} = -2[LL(\beta_1) - LL(\beta_2)] \quad (7)$$

in which $LL(\beta_1)$ and $LL(\beta_2)$ represent the log-likelihoods at convergence of the two competing models. The χ^2 statistic is indicative of the chi-square distribution, with degrees of freedom equal to the difference between the numbers of variables in models 1 and 2 (Washington et al., 2020).

The model predictive accuracy is evaluated using 20% of the observations not used for the model development (Saheli and Effati, 2021; Sheikholeslami Bondarabadi, and Asadamraji, 2020). Applied validation indicators include the mean absolute deviation (MAD) and the root mean squared error (RMSE):

$$MAD = \frac{1}{N} \sum_{t=1}^N |y_{predicted} - y_{observed}| \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_{predicted} - y_{observed})^2} \quad (9)$$

Here, $y_{observed}$ and $y_{predicted}$ represent the observed and predicted values, respectively, while N indicates the number of observations in the validation process. The selection among other models is based on the model exhibiting the lowest RMSE and MAD.

Derived from the estimated parameters, the average marginal effects and elasticity values can be utilized to interpret how the explanatory independent variables affect crash frequency. The average marginal effect of an explanatory variable signifies the impact of a one-unit change in that variable on the predicted crash frequency, holding the values of other explanatory variables constant at their mean value. The marginal effect of explanatory variable l is expressed by the following equation:

$$ME_{X_{il}}^{\lambda_i} = \frac{\partial \lambda_i}{\partial X_{il}} = \beta_l EXP(\beta x) \quad (10)$$

in which x and β represent the vectors of explanatory variables and their corresponding parameter estimates, respectively.

The average elasticity value indicates the percentage change in crash frequency for a 1% change in a continuous explanatory variable. The following shows the calculation of elasticity for the average of explanatory variable l :

$$E_{X_{il}}^{\lambda_i} = \left(\frac{\partial \lambda_i}{\lambda_i}\right) \left(\frac{X_{il}}{\partial X_{il}}\right) = \beta_l X_{il} \quad (11)$$

The value of explanatory variable l for the i th observation is represented by X_{il} , while β_l signifies the estimated parameter for explanatory variable l .

4. Data

The current research aims to develop crash frequency models in the marginal regions

surrounding cities. To achieve this objective, the study utilized rural crash data from highways leading to Isfahan, i.e., those where travel originates from or terminates at Isfahan. Due to its position as the second largest city in Iran and its central location within the country, Isfahan experiences significant traffic volume and crash frequency on the highways leading to the city. An assessment of the road network reveals six highways connecting to the city of Isfahan: Isfahan-Zarrinshahr, Isfahan-Skahreza, Isfahan-Naein, Isfahan-Airport, Isfahan-Shahinshahr, and Isfahan-Murchehkhort.

The models introduced in a study conducted by Shamanian Esfahani, Afandizadeh, and Naderan (2022) enable the determination of the boundary of the influence area. The values yielded by the models for the influence area of the highways leading to Isfahan ranged between 10 km and 20 km.

The rural crash database in the marginal areas of Isfahan was established by delineating the extent of influence for each highway. The database comprised 2613 crashes that had taken place during a three-year period from 2016 to 2019. The crash data was obtained from the highway police department.

Following the compilation of the rural crash database within Isfahan's influence area, the distance of each crash from the city was established. Initially, shapefiles of the city, highways leading to the city, and crashes in its marginal areas were imported into the ArcGIS software to facilitate this process. Subsequently, the distance of each crash occurrence from the city of Isfahan was calculated along the studied highways.

Creating crash frequency estimation models necessitated segmenting the studied highways. The static method (fixed-length) was employed to segment studied roads, as in some previous crash frequency modelling studies (Cafiso et al., 2013; Cenek et al., 1997; Shankar et al., 1995). Segmentation was carried out based on 5-km intervals to ensure homogeneity in road geometry and traffic flow characteristics,

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

including speed limit, traffic volume, shoulder type, shoulder width, and lane count.

For each highway, segment number 1 denoted the closest segment to the city, with subsequent segments numbered sequentially based on their increasing distance from the city. For example, the Isfahan-Zarrinshahr highway, with a length of 20 km, was divided into four 5-km segments. Segment 1, situated nearest to Isfahan, terminated at a distance of 5 km from the city boundary. Subsequent segments were numbered accordingly. In total, the studied highways were segmented into nineteen 5-km segments. Following the segmentation of highways and accounting for the distance of each crash occurrence from the city, the number of crashes in each segment was determined. Table 1 presents the data regarding the number of crashes in each segment.

The additional data collected encompassed various roadway/geometric attributes (type of highway, speed limit, count of controlled and uncontrolled access points, average spacing between access points, number of speed cameras, availability of rest area amenities, lane count, shoulder type, shoulder width, and longitudinal gradient), traffic-related features (traffic volume, percentage of heavy vehicles, ratio of local to non-local vehicles, proportion of passing vehicles during daytime and nighttime, average vehicle speed, percentage of vehicles exceeding speed limits, and percentage of vehicles breaching longitudinal distance limits), characteristics of the built environment/land use (length of adjacent land use), and meteorological/ environmental factors (total monthly precipitation).

The roadway/geometric attributes of the examined highways were provided by the Road Maintenance and Transportation Organization and the Road and Urban Development Bureau of Isfahan Province. The length of adjacent land use was determined by analyzing aerial images and maps of the studied roads, complemented by field observations to ensure accuracy.

Additionally, traffic data was collected utilizing information from loop detectors installed along the studied highways. Data regarding the monthly rainfall on the studied highways was acquired from the weather stations located in Isfahan Province.

Based on the monthly crash data for each 5-km segment of the studied highways, along with the aforementioned traffic, geometric, and environmental information, the final table was generated, comprising 540 rows, equivalent to 80% of the observations. The remaining observations were utilized for assessing the accuracy of the models. Each row in the table represented a specific observation corresponding to one segment of the studied highways. Table 2 presents the descriptive analysis of the continuous variables, while Table 3 displays the descriptive analysis of the categorical variables.

In Table 2, the variable "number of crashes" denotes the count of monthly crashes in the studied segments. Additionally, the variables pertaining to the traffic characteristics of the studied segments, along with the precipitation variable, are also reported on a monthly basis. The "Isfahan vehicles traffic volume" denotes the count of vehicles registered with license plates from Isfahan. Drivers of these vehicles can be regarded as local and familiar with the route. The "proportion of speed-limit violations" represents the percentage of detected vehicles that exceed the designated highway speed limit. The "longitudinal distance limit violation" occurs when the minimum following distance of 2 seconds from the preceding vehicle is not maintained. The proportion of nighttime traffic flow represents the percentage of total traffic passing through the studied segment during nighttime. One of the variables listed in Table 3 is "access to rest area facilities." If a segment is situated on a highway with access to rest area facilities for drivers, this value is represented as one; otherwise, it is zero.

Table 1. The crash frequency of each studied highways segment

Highway	Segment			
	1	2	3	4
Isfahan-Zarinshahr	254	222	244	215
Isfahan-Shahreza	235	212	148	-
Isfahan-Naeen	85	77	72	-
Isfahan-Airport	104	101	-	-
Isfahan-Shahinshahr	82	88	68	50
Isfahan-Mourchekhort	114	138	104	-

Table 2. Descriptive statistics of continues variables

Variable	Mean	Min.	Max.	Std. Dev.
Number of crashes	3.8	1.0	14.0	2.5
Traffic volume (10,000 veh)	71.1	22.1	154.9	42.8
Isfahan vehicles traffic volume (10,000 veh)	51.3	13.4	126.1	34.5
Other vehicles traffic volume (10,000 veh)	19.7	7.5	43.2	9.0
Proportion of heavy vehicles (%)	13.4	2.7	26.2	6.9
Average speed (km/hr)	89.9	83.0	97.0	4.8
Proportion of speed-limit violation (%)	10.9	2.3	26.1	5.8
Proportion of longitudinal distance-limit violation (%)	13.6	4.0	27.6	6.9
Proportion of night traffic volume (%)	30.3	24.7	34.7	3.4
Rainfall (mm)	7.4	0.0	28.2	7.3
Number of Lanes	2.4	2.0	3.0	0.5
Shoulder width (m)	3.1	2.0	3.5	0.5
Speed limit (km/hr)	114.2	110.0	120.0	4.9
Longitudinal grade (%)	0.9	0.4	1.7	0.4
Number of speed cameras	1.8	0.0	4.0	1.2
Length of adjacent land use	2231.6	500.0	4800.0	1330.1
Number of access points	10.2	4.0	18.0	4.4
Average distance between access points	523.4	154.0	784.0	210.2
Number of controlled access points	4.6	2.0	10.0	1.8
Number of uncontrolled access points	5.6	0.0	15.0	3.9

Table 3. Descriptive statistics of categorical variables

Variable	Category	Frequency
Indicator variable for freeway	1	216
	0	324
Indicator variable for Expressway	1	324
	0	216
Indicator variable for paved shoulder	1	288
	0	252
Indicator variable for access to rest area facilities	1	324
	0	216

5. Modelling Results and Discussion

The models developed in this study are RENB and RPNB. Table 4 displays the outcomes of the model estimation. Variables that were not

International Journal of Transportation Engineering,
Vol. 12/ No. 2/ (46) Autumn 2024

statistically significant at the 0.05 level were removed based on the results of the correlation analysis using the Spearman coefficient and a backward stepwise regression approach. The simulation-based maximum likelihood method that maximizes the simulated log-

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

likelihood function was used to estimate the RPNM model. In order to improve the model's accuracy and ensure reliable estimations, 200 Halton draws were utilized, as recommended in previous studies (Russo et al., 2014; Anastasopoulos et al., 2012). The examination of various model parameter distributions, including log-normal, normal, triangular, and uniform ones, was conducted using the RPNB model. The statistical analysis revealed that all parameters exhibited the best fit with the normal distribution, consistent with findings from prior research (Mahmud et al., 2019; Chen and Tarko, 2014).

Out of the models, a total of seven parameters were identified as significant, with six of them having variables displaying significant variances. Consequently, these variables were considered random in the RPNB model. Variables such as intercept, Isfahan vehicles traffic volume, other vehicles traffic volume, the number of controlled access points, the number of uncontrolled access points, and the indicator for access to rest area facilities were estimated as random variables. It was observed that all estimated coefficients in the models exhibited reasonable signs.

Table 4. Model estimation results

Description	RENB		RPNB	
	Estimate	z-stat	Estimate	z-stat
Intercept	1.618	12.94	1.627	13.15
Standard deviation of distribution	0.315	63.76	0.191	38.46
Isfahan vehicles traffic volume	0.013	5.96	0.013	6.12
Standard deviation of distribution			0.019	14.48
Other vehicles traffic volume	-0.049	-7.54	-0.051	-7.93
Standard deviation of distribution			0.034	27.09
Proportion of longitudinal distance-limit violation	0.014	2.16	0.015	2.28
Number of controlled access points	-0.031	-2.50	-0.030	-2.41
Standard deviation of distribution			0.025	7.63
Number of uncontrolled access points	0.027	2.76	0.025	2.87
Standard deviation of distribution			0.022	5.74
Indicator variable for access to rest area facilities				
0 (Reference group)				
1	-0.494	-5.85	-0.523	-6.19
Standard deviation of distribution			0.288	7.14
Dispersion parameter	0.278	2.22	0.282	2.16
Goodness-of-fit-measures				
Log-likelihood at convergence	1566.113		-1560.427	
Log-likelihood with constant only	-2868.848		-2868.848	
McFadden ρ^2	0.454		0.456	
AIC	3159.3		3148.9	
BIC	3216.0		3204.8	
Forecasting accuracy				
MAD	0.64		0.62	
RMSE	0.82		0.81	
Likelihood ratio test				
Number of parameters	9		14	
Degree of freedom		5		
Level of confidence		95%		
Computed chi-square		11.37		

Description	RENB		RPNB	
	Estimate	z-stat	Estimate	z-stat
Statistically superior model	RPNB			

To assess goodness-of-fit, the models were compared using measures such as AIC, McFadden ρ^2 , and BIC. For the RENB model, the McFadden ρ^2 stands at 0.454, while in the RPNB model, it slightly increases to 0.456. Regarding AIC, the RENB model records 3159.3, whereas the RPNB model shows a slightly lower value of 3148.9. Comparable trends are noted in the changes in BIC.

Both MAD and RMSE were utilized to assess the models' forecasting accuracy. Upon comparing the RENB and RPNB models, no significant difference in forecasting accuracy was observed. However, assessment measures slightly improved for the RPNB model compared to the RENB model. In the RPNB model, the MAD and RMSE values were lower by 3.23% and 1.23%, respectively, compared to the RENB model.

The analysis of the developed models demonstrated the RPNB model's superiority over the RENB model within a 0.05 confidence interval. However, upon comparing the calculated χ^2 value (11.37 with five degrees of freedom) to the critical value of chi-square (11.07), the performance difference between the two models was not deemed significant.

Although the results showcased higher values of the log-likelihood at convergence and McFadden ρ^2 in the RPNB model compared to the RENB model, suggesting its superior performance in terms of goodness-of-fit, the marginal difference in measures between the RENB and RPNB models indicated their nearly equivalent performance. The enhancement in the measures of McFadden ρ^2 , AIC, and BIC in the RPNB model, as opposed to the RENB model, was approximately 0.44%, 0.33%, and 0.35%, respectively. Conversely, the results of the likelihood ratio test confirmed the marginal difference in performance between the two models. This observation is consistent with the

findings of a parallel study carried out by Chen and Tarko (2014). Nevertheless, the higher standard deviation of the distribution for the intercept in the RENB model, compared to that of the RPNB model, suggests that the intercept of the RENB model exhibits a greater degree of unobserved heterogeneity than the corresponding parameter in the RPNB model. Another part of heterogeneity in the RPNP model was explained by other parameters. This could be the main reason for the small difference in the likelihood function of the two models and their almost similar performances. Table 5 lists the values of marginal effects and elasticity for the RENB and RPNB models separately for each of the descriptive variables. As expected, the results indicated a small difference in the marginal effects of the explanatory variables in the RPNB model compared to the RENB model.

Most road safety studies have discussed the effect of road traffic volume on crash frequency (Mousavi et al., 2019; Yu et al., 2019; Ma et al., 2017; Anastasopoulos and Mannering, 2009). This study showed that not only is there a relationship between the traffic volume and crash frequency, but also the combination of local and non-local vehicles and the proportion of each group of the total road traffic volume affect the crash frequency. According to the modeling results, the number of crashes increased with the rise of local vehicle percentage in the marginal areas around the cities. On the other hand, the rise in the proportion of non-local vehicles in marginal areas reduced the probability of crash occurrence. This can be due to the riskier behaviors of local drivers compared to the non-local ones due to their familiarity with the route in the marginal areas of the cities. This was approved by the positive coefficient of the Isfahan vehicles' traffic volume variable,

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

besides the negative coefficient of the other vehicle's traffic volume variable. According to the RPNB model, the traffic volume of the local vehicles had a normal distribution with a mean of 0.013 and a standard deviation of 0.019. The rise in the traffic volume of local vehicles increased the crash frequency in 75.3% of the observations while reducing it in 24.7% of the observations.

Moreover, the other vehicle's traffic volume variable followed a normal distribution with a mean of -0.051 and a standard deviation of 0.034. The outcomes revealed that the increased traffic volume of non-local vehicles resulted in a lower crash frequency in 93.3% of observations and a higher crash frequency in 6.7% of the observations. The effect of the proportion of local and non-local vehicles on crashes has been approved by the results of a study by Yu et al. (2019).

Some road safety studies have confirmed the relationship between speed limit violations and rural crash frequency (Afandizadeh and Hassanpour, 2020; Ehsani Sohi, Dashtestaninejad, and Khademi, 2019). However, the present study showed that in the marginal areas of cities, the relationship between the longitudinal distance limit violation and crash frequency should be taken into account. Based on the results of the RPNB model, with a 1% rise in the proportion of the longitudinal distance limit violation, the monthly crash frequency of the segments increased by 0.2%, on average. The increased proportion of the longitudinal distance limit violation indicated the decreased distance between the vehicles. Given the different land uses along the segments located in the marginal areas of cities and the large number of access roads in these areas, the probability of a reduction in the speed of vehicles due to the merge and diverge maneuvers is high (Ahmadinejad, Afandizadeh Zargari, and Jalalkamali, 2018). As a result, in the case of violation of the longitudinal distance limit by

vehicles, the probability of collision and crash occurrence increases.

The modeling results showed that the rise in the proportion of uncontrolled access points increases the crash frequency in the marginal areas around cities. According to the RPNB model, the number of uncontrolled access points had a normal distribution with a mean of 0.025 and a standard deviation of 0.022. In 87.2% of the observations, the crash frequency increased with the number of uncontrolled access points. The relationship between this variable and crash frequency for other observations (12.8%) was inverse. On average, a 1% rise in the number of uncontrolled access points increased the monthly crash frequency of each segment by 0.14%.

On the other hand, an increased proportion of controlled access points reduced the crash frequency. The number of controlled access points followed a normal distribution with a mean of -0.030 and a standard deviation of 0.025. As the number of controlled access points increased, the number of crashes declined in 88.5% of the observations while growing in the remaining (11.5%) observations. The relationship between the control of access points and the rural crash frequency found in this study agrees with the result of the study by Mohammadnazar et al. (2021a). However, some other studies mentioned that the increased number of access points, regardless of how they are controlled, resulted in a higher crash frequency (Ehsani Sohi, Dashtestaninejad, and Khademi, 2021; Hou, Tarko, and Meng, 2018; Dinu and Veeraragavan, 2011).

The obtained results revealed the positive effect of the presence of rest area facilities on the safety of rural highways in the marginal zones of cities. The calculated value of the marginal effect for the RPNB model indicated that the access to rest area facilities reduced the crash frequency in the segments in the marginal areas of cities by 1.68 crashes/month on average. For drivers who travel long distances to reach a city, resting in a rest area facility can reduce fatigue

and prevent needless hurry to reach the destination, which is in turn a reason for the increased crashes.

Table 5. Average marginal effects and elasticity values for RENB model and RPNB model

Variable	RENB		RPNB	
	Marginal Effects	Elasticity	Marginal Effects	Elasticity
Isfahan vehicles traffic volume	0.04	0.67	0.04	0.67
Other vehicles traffic volume	-0.16	-1.00	-0.16	-1.01
Proportion of longitudinal distance-limit violation	0.05	0.19	0.05	0.20
Number of controlled access points	-0.10	-0.14	-0.10	-0.14
Number of uncontrolled access points	0.09	0.15	0.08	0.14
Indicator variable for access to rest area	-1.65	-	-1.68	-

6. Conclusion

This study was performed to identify the parameters affecting the crash frequency in the marginal areas of cities. These areas have a considerable proportion of rural crashes in Iran. In this study, the rural crash data of six highways leading to Isfahan were used. The approach of random-parameter was employed to develop crash frequency models for these parts of rural highways. Comparison of the performance of random-parameter models considering different assessment measures, especially likelihood ratio, showed the slight superiority of the RPNB models over the RENB ones.

The estimation results indicated that crash frequency in the marginal areas of cities is affected by the traffic and environmental characteristics of highways. Increased traffic volume of local vehicles, the proportion of longitudinal distance limit violations, and the number of uncontrolled access points increase the crash frequency around the cities. On the other hand, the increased volume of non-local vehicles, increased number of controlled access points, and access to rest area facilities result in a lower frequency of crashes in these areas. According to the obtained results, not only is there a relationship between the traffic volume and crash frequency but also the combination of local and non-local vehicles and the proportion of each group of the total road traffic volume

affect the crash occurrence probability. Given the presence of land uses along the highways and increased access roads in segments located in the marginal areas of cities compared to the basic segments of the highways, and consequently, increased merge and diverge maneuvers in these segments, the violation of longitudinal distance limit increases the probability of crash occurrence. The positive coefficient of the proportion of longitudinal distance limit violation proved this finding. The effect of access points controlling on the crash frequency is another finding of this study, which can help decision-makers implement solutions effective in reducing the rural crashes around cities. The effect of access to rest areas on reducing the crash frequency indicates the necessity of further attention to developing such facilities on rural highways.

The findings of this study give a better understanding of the factors effective in the occurrence of crashes in the marginal areas of cities and the implementation of the related safety measures according to their characteristics. Implementation of efficient measures in such areas not only reduces the rural crashes and the losses caused by them but also reduces the costs required to make highways safer, given the numerous financial limitations and problems in developing countries.

As a limitation of crash modelling studies, it is too difficult to measure all important

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

parameters affecting crash frequency estimation models. So, some factors were not measured and used in this study, such as road pavement, horizontal and vertical alignment of the road, type of adjacent land use, etc. Future research is needed to evaluate the effects of these factors on marginal areas' rural crash frequencies. Accessing more detailed reports about road conditions prepared by the RMTO organization could be a proper method to collect these factors data.

Pedestrian or motorcycle crash frequency modelling in marginal areas around cities is a thoughtful issue that could be addressed in future studies. Also, the comparison of effective parameters on crashes in marginal and basic segments of roads is another suggestion for future studies.

7. References

– Abdel-Aty, M. A. and Radwan, A. E. (2000) “Modeling traffic accident occurrence and involvement”, *Accident Analysis & Prevention*, 32(5), 633-642.

[https://doi.org/10.1016/S0001-4575\(99\)00094-9](https://doi.org/10.1016/S0001-4575(99)00094-9)

– Abdel-Aty, M., Devarasetty, P. C., and Pande, A. (2009) “Safety evaluation of multilane arterials in Florida”, *Accident Analysis & Prevention*, 41(4), 777-788.

<https://doi.org/10.1016/j.aap.2009.03.015>

– Afandizadeh, S. and Hassanpour, S. (2020) “Evaluating the Effect of Roadway and Development Factors on the Rural Road Safety Risk Index”, *Advances in Civil Engineering*, 2020.

<https://doi.org/10.1155/2020/7820565>

– Afghari, A. P., Washington, S., Haque, M. M., and Li, Z. (2018) “A comprehensive joint econometric model of motor vehicle crashes arising from multiple sources of risk”, *Analytic methods in accident research*, 18, 1-14.

<https://doi.org/10.1016/j.amar.2018.03.002>

– Ahmadinejad, M., Afandizadeh Zargari, S., and Jalalkamali, R. (2018) “Are deceleration numbers a suitable index for road safety?”, In *Proceedings of the Institution of Civil Engineers-Transport* (Vol. 171, No. 5, pp. 247-252). Thomas Telford Ltd.

<https://doi.org/10.1680/jtran.16.00117>

– Ahmed, M., Huang, H., Abdel-Aty, M., and Guevara, B. (2011) “Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway”, *Accident Analysis & Prevention*, 43(4), 1581-1589.

<https://doi.org/10.1016/j.aap.2011.03.021>

– Alrejail, A., Moomen, M., and Ksaibati, K. (2021) “Evaluating the effectiveness of law enforcement in reducing truck crashes for a rural mountainous freeway in Wyoming”, *Transportation Letters*, 1-11.

<https://doi.org/10.1080/19427867.2021.1949089>

– Anastasopoulos, P. C. and Mannering, F. L. (2009) “A note on modeling vehicle accident frequencies with random-parameters count models”, *Accident Analysis & Prevention*, 41(1), 153-159.

<https://doi.org/10.1016/j.aap.2008.10.005>

– Anastasopoulos, P. C., Mannering, F. L., Shankar, V. N., and Haddock, J. E. (2012) “A study of factors affecting highway accident rates using the random-parameters tobit model. Accident”, *Analysis & Prevention*, 45, 628-633.

<https://doi.org/10.1016/j.aap.2011.09.015>

– Ayati, E., and Abbasi, E. (2011) “Investigation on the role of traffic volume in accidents on urban highways”, *Journal of safety research*, 42(3), 209-214.

<https://doi.org/10.1016/j.jsr.2011.03.006>

– Boroujerdian, A.M., Saffarzadeh, M. and

- Abolhasannejad, V. (2010) "Developing a model for prioritising high crash road segments", *Proceedings of the Institution of Civil Engineers-Transport*, Vol. 163, No. 1, pp. 19-29.
<https://doi.org/10.1680/tran.2010.163.1.19>
- Buddhavarapu, P., Banerjee, A., and Prozzi, J. A. (2013) "Influence of pavement condition on horizontal curve safety", *Accident Analysis & Prevention*, 52, 9-18.
<https://doi.org/10.1016/j.aap.2012.12.010>
- Cai, Q., Abdel-Aty, M., Lee, J., Wang, L., and Wang, X. (2018) "Developing a grouped random-parameters multivariate spatial model to explore zonal effects for segment and intersection crash modeling", *Analytic methods in accident research*, 19, 1-15.
<https://doi.org/10.1016/j.amar.2018.05.001>
- Cafiso, S., D'Agostino, C., & Persaud, B. (2013). Investigating Influence of Segmentation in Estimating Safety Performance Functions for Roadway Sections (No. 13-4372).
- Cenek, P.D., Davies, R.B., McLarin, M.W., Griffith-Jones, G., & Locke, N.J. (1997). Roadenvironment and traffic crashes. In: *Transfund New Zealand Research Report 79*, Wellington.
- Chen, E., and Tarko, A. P. (2014) "Modeling safety of highway work zones with random-parameters and random effects models", *Analytic methods in accident research*, 1, 86-95.
<https://doi.org/10.1016/j.amar.2013.10.003>
- Dinu, R. R., and Veeraragavan, A. (2011) "Random-parameter models for accident prediction on two-lane undivided highways in India", *Journal of safety research*, 42(1), 39-42.
<https://doi.org/10.1016/j.jsr.2010.11.007>
- Dong, C., Nambisan, S. S., Richards, S. H., and Ma, Z. (2015) "Assessment of the effects of highway geometric design features on the frequency of truck involved crashes using bivariate regression", *Transportation Research Part A: Policy and Practice*, 75, 30-41.
<https://doi.org/10.1016/j.tra.2015.03.007>
- Effati, M., Rajabi, M.A. and Samadzadegan, F. (2012) "Developing a novel method for road hazardous segment identification based on fuzzy reasoning and GIS", *Journal of Transportation Technologies*, Vol. 2, No. 1, pp. 32-40.
<http://dx.doi.org/10.4236/jtts.2012.21004>
- Ehsani Sohi, M., Dashtestaninejad, H. and Khademi, E. (2019) "Effects of Roadway and Traffic Characteristics on Accidents Frequency at City Entrance Zone", *International Journal of Transportation Engineering*, 7(2), pp.139-152.
<http://dx.doi.org/10.22119/IJTE.2018.130189.1404>
- Elyasi, M. R., Saffarzadeh, M. and Boroujerdian, A. M. (2016) "A novel dynamic segmentation model for identification and prioritization of black spots based on the pattern of potential for safety improvement", *Transportation Research Part A*, Vol. 91, pp. 346-357.
<https://doi.org/10.1016/j.tra.2016.06.020>
- Elyasi, M.R., Saffarzadeh, M. and Boroujerdian, A.M. (2018) "Assessing the interrelations of traffic collisions' risk factors", *Proceedings of the Institution of Civil Engineers-Transport*, Vol. 171, No. 6, pp. 309-318.
<https://doi.org/10.1680/jtran.16.00070>
- Faezi, S. F., and Elyasi, M. R. (2020) "Predict road accidents resulted from carelessness using negative binomial regression model (A case study of rural highways in Hamadan

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

Province)”, *Journal of Transportation Research*. Available:http://www.trijournal.ir/article_119190.html?lang=en

– Fu, R., Guo, Y., Yuan, W., Feng, H., and Ma, Y. (2011) “The correlation between gradients of descending roads and accident rates”, *Safety science*, 49(3), 416-423.

<https://doi.org/10.1016/j.ssci.2010.10.006>

– Garach, L., de Oña, J., López, G., and Baena, L. (2016) “Development of safety performance functions for Spanish two-lane rural highways on flat terrain”, *Accident Analysis & Prevention*, 95, 250-265.

<https://doi.org/10.1016/j.aap.2016.07.021>

– Gong, H., Wang, F., Zhou, B. B., and Dent, S. (2020) “Application of random effects negative binomial model with clustered dataset for vehicle crash frequency analysis”, *International journal of transportation science and technology*, 9(3), 183-194.

<https://doi.org/10.1016/j.ijtst.2020.03.010>

– Han, C., Huang, H., Lee, J., and Wang, J. (2018) “Investigating varying effect of road-level factors on crash frequency across regions: A Bayesian hierarchical random-parameter modeling approach”, *Analytic methods in accident research*, 20, 81-91.

<https://doi.org/10.1016/j.amar.2018.10.002>

– Hosseinlou, M. H., Mahdavi, A., and Nooghabi, M. J. (2018) “Validation of the influencing factors associated with traffic violations and crashes on freeways of developing countries: A case study of Iran”, *Accident Analysis & Prevention*, 121, 358-366.

<https://doi.org/10.1016/j.aap.2018.06.009>

– Hou, Q., Tarko, A. P., and Meng, X. (2018) “Investigating factors of crash frequency with random effects and random-parameters models: new insights from Chinese freeway study”,

Accident Analysis & Prevention, 120, 1-12.

<https://doi.org/10.1016/j.aap.2018.07.010>

– Imaninasab, R., Sekhavati, E., and Hajihoseinloo, M. (2016) “Modeling crash frequency of heavy vehicles in rural freeways”, *Journal of Traffic and Logistics Engineering* Vol, 4(2).

<http://dx.doi.org/10.18178/jtle.4.2.98-102>

– Iran Road Maintenance and Transportation Organization (RMTO). (2022) “Road Statistical Yearbook”, Available:

<http://www.rmto.ir/Pages/Home.aspx>

– Khoda Bakhshi, A., and Ahmed, M. M. (2021) “Coping with endogeneity and unobserved heterogeneity in real-time clustering critical crash occurrences nested within weather and road surface conditions”, *International journal of injury control and safety promotion*, 28(2), 208-221.

<https://doi.org/10.1080/17457300.2021.1907595>

– Kim, K., Brunner, I. M., and Yamashita, E. Y. (2006) “Influence of land use, population, employment, and economic activity on accidents”, *Transportation research record*, 1953(1), 56-64.

<https://doi.org/10.1177/0361198106195300107>

– Lord, D., Washington, S.P., and Ivan, J.N. (2007) “Further notes on the application of zero inflated models in highway safety”, *Accident Analysis and Prevention*, 39 (1), 53–57.

<https://doi.org/10.1016/j.aap.2006.06.004>

– Lord, D., and Mannering, F. (2010) “The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives”, *Transportation research part A: policy and practice*, 44(5), 291-305.

<https://doi.org/10.1016/j.tra.2010.02.001>

- Ma, Z., Zhang, H., Steven, I., Chien, J., Wang, J., and Dong, C. (2017) “Predicting expressway crash frequency using a random effect negative binomial model: A case study in China”, *Accident Analysis & Prevention*, 98, 214-222.
<https://doi.org/10.1016/j.aap.2016.10.012>
- Mahmud, S. S., Ferreira, L., Hoque, M. S., and Tavassoli, A. (2019) “Micro-level safety risk assessment model for a two-lane heterogeneous traffic environment in a developing country: A comparative crash probability modeling approach”, *Journal of Safety Research*, 69, 125-134.
<https://doi.org/10.1016/j.jsr.2019.03.008>
- Malyshkina, N. V., and Mannering, F. L. (2010) “Zero-state Markov switching count-data models: An empirical assessment”, *Accident Analysis & Prevention*, 42(1), 122-130.
<https://doi.org/10.1016/j.aap.2009.07.012>
- Mannering, F. L., and Bhat, C. R. (2014) “Analytic methods in accident research: Methodological frontier and future directions”, *Analytic methods in accident research*, 1, 1-22.
<https://doi.org/10.1016/j.amar.2013.09.001>
- Mannering, F. L., Shankar, V., and Bhat, C. R. (2016) “Unobserved heterogeneity and the statistical analysis of highway accident data”, *Analytic methods in accident research*, 11, 1-16.
<https://doi.org/10.1016/j.amar.2016.04.001>
- Mannering, F., Bhat, C. R., Shankar, V., and Abdel-Aty, M. (2020) “Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis”, *Analytic methods in accident research*, 25, 100113.
<https://doi.org/10.1016/j.amar.2020.100113>
- Mohammadnazar, A., Arvin, R., and Khattak, A. J. (2021) “Classifying travelers' driving style using basic safety messages generated by connected vehicles: Application of unsupervised machine learning”, *Transportation research part C: emerging technologies*, 122, 102917.
<https://doi.org/10.1016/j.trc.2020.102917>
- Mohammadnazar, A., Mahdinia, I., Ahmad, N., Khattak, A. J., and Liu, J. (2021a) “Understanding how relationships between crash frequency and correlates vary for multilane rural highways: Estimating geographically and temporally weighted regression models”, *Accident Analysis & Prevention*, 157, 106146.
<https://doi.org/10.1016/j.aap.2021.106146>
- Mohaymany, A.S., Shahri, M. and Mirbagheri, B. (2013) “GIS-based method for detecting high-crash-risk road segments using network kernel density estimation”, *Geo-spatial Information Science*, 16(2), pp.113-119.
<https://doi.org/10.1080/10095020.2013.766396>
- Moomen, M., Rezapour, M., Raja, M., and Ksaibati, K. (2020) “Predicting downgrade crash frequency with the random parameters negative binomial model: Insights into the impacts of geometric variables on downgrade crashes in Wyoming”, *IATSS research*, 44(2), 94-102.
<https://doi.org/10.1016/j.iatssr.2019.09.002>
- Mousavi, S. M., Marzoughi, H., Parr, S. A., Wolshon, B., and Pande, A. (2019) “A mixed crash frequency estimation model for interrupted flow segments”, In *International Conference on Transportation and Development 2019: Smarter and Safer Mobility and Cities* (pp. 72-83). Reston, VA: American Society of Civil Engineers. Available:
<https://ascelibrary.org/doi/10.1061/9780784482575.008>

Determination of Factors Affecting Rural Crash Frequency at Marginal Areas around Cities

- Naik, B., Tung, L. W., Zhao, S., and Khattak, A. J. (2016) “Weather impacts on single-vehicle truck crash injury severity”, *Journal of safety research*, 58, 57-65.
<https://doi.org/10.1016/j.jsr.2016.06.005>
- Nassiri, H., Najafi, P., and Mohamadian Amiri, A. (2014) “Prediction of roadway accident frequencies: Count regressions versus machine learning models”, *Scientia Iranica*, 21(2), 263-275. Available:
http://scientiairanica.sharif.edu/article_1630.html
- Park, J., Abdel-Aty, M., and Wang, J. H. (2017) “ime series trends of the safety effects of pavement resurfacing”, *Accident Analysis & Prevention*, 101, 78-86.
<https://doi.org/10.1016/j.aap.2017.02.006>
- Rusli, R., Haque, M. M., King, M., and Voon, W. S. (2017) “Single-vehicle crashes along rural mountainous highways in Malaysia: an application of random parameters negative binomial model”, *Accident Analysis & Prevention*, 102, 153-164.
<https://doi.org/10.1016/j.aap.2017.03.002>
- Russo, B. J., Savolainen, P. T., Schneider IV, W. H., and Anastopoulos, P. C. (2014) “Comparison of factors affecting injury severity in angle collisions by fault status using a random-parameters bivariate ordered probit model”, *Analytic methods in accident research*, 2, 21-29.
<https://doi.org/10.1016/j.amar.2014.03.001>
- Saeed, T. U., Hall, T., Baroud, H., and 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.
<https://doi.org/10.1016/j.amar.2019.100101>
- Saheli, M.V. and Effati, M. (2021) “Segment-Based Count Regression Geospatial Modeling of the Effect of Roadside Land Uses on Pedestrian Crash Frequency in Rural Roads”, *International Journal of Intelligent Transportation Systems Research*, Vol. 19, No. 2, pp. 347-365.
- Sajed, Y., Shafabakhsh, G. and Bagheri, M. (2019) “Hotspot Location Identification Using Accident Data, Traffic and Geometric Characteristics”, *Engineering Journal*, 23(6), pp.191-207.
<https://doi.org/10.4186/ej.2019.23.6.191>
- Shafabakhsh, G.A., Famili, A. and Akbari, M. (2016) “Spatial analysis of data frequency and severity of rural accidents”, *Transportation Letters*, pp.1-8.
<https://doi.org/10.1080/19427867.2016.1138605>
- Shamanian Esfahani, H., Afandizadeh, S., Naderan, A. (2022) “Investigating the Effect of Marginal Areas around the Cities on Rural Road Accidents in Iran Using Linear and Logistic Regression Approaches”, *International Journal of Transportation Engineering*, 9(3), 693-711.
doi: 10.22119/ijte.2021.284119.1567
- Shankar, V., Mannering, F., & Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention*, 27(3), 371-389.
- Shaon, M. R. R., Qin, X., Shirazi, M., Lord, D., and Geedipally, S. R. (2018) “Developing a Random parameters negative binomial-Lindley Model to analyze highly over-dispersed crash count data”, *Analytic methods in accident research*, 18, 33-44.
<https://doi.org/10.1016/j.amar.2018.04.002>
- Sheikholeslami, S., Bondarabadi, M.A. and

Asadamraji, M. (2020) "A rural road accident probability model based on single vehicle hazard properties including hazard color and mobility: a driving simulator study", *Journal of advanced transportation*.

– W.H. (2012) "Greene, *Econometric Analysis*", Seventh, Prentice Hall, Boston, MA, 2012.

– Washington, S., Karlaftis, M., Mannering, F., and Anastasopoulos, P. (2020) "Statistical and econometric methods for transportation data analysis", Chapman and Hall/CRC.
<https://doi.org/10.1201/9780429244018>

– Wen, X., Xie, Y., Wu, L., and Jiang, L. (2021) "Quantifying and comparing the effects of key risk factors on various types of roadway segment crashes with LightGBM and SHAP", *Accident Analysis & Prevention*, 159, 106261.
<https://doi.org/10.1016/j.aap.2021.106261>

– World Health Organization (WHO). 2018 "Global Status Report on Road safety", Available:
<https://www.who.int/publications/i/item/global>

– status-report-on-road-safety-2018.

– Wu, P., Song, L., and Meng, X. (2021) "Influence of built environment and roadway characteristics on the frequency of vehicle crashes caused by driver inattention: a comparison between rural roads and urban roads", *Journal of safety research*, 79, 199-210.
<https://doi.org/10.1016/j.jsr.2021.09.001>

– Yan, Y., Zhang, Y., Yang, X., Hu, J., Tang, J., and Guo, Z. (2020) "Crash prediction based on random effect negative binomial model considering data heterogeneity", *Physica A: Statistical Mechanics and Its Applications*, 547, 123858.
<https://doi.org/10.1016/j.physa.2019.123858>

– Yu, R., Wang, Y., Quddus, M., and Li, J. (2019) "A marginalized random effects hurdle negative binomial model for analyzing refined-scale crash frequency data", *Analytic methods in accident research*, 22, 100092.

<https://doi.org/10.1016/j.amar.2019.100092>

– Zou, Y., and Tarko, A. P. (2016) "An insight into the performance of road barriers—redistribution of barrier-relevant crashes", *Accident Analysis & Prevention*, 96, 152-161.
<https://doi.org/10.1016/j.aap.2016.07.022>