TRANSPORTES
revistatransportes.org.br
anpet
associação nacional de pesquisa e ensino em transportes
anpet.org.br

An analysis of the injury severity of pedestrians in Brazil using random parameters logit models

Análise da gravidade de lesões de pedestres no Brasil utilizando modelos logit de parâmetros aleatórios

Mateus Nogueira Silva[1], **Flávio José Craveiro Cunto**[2], **Marcos José Timbó Lima Gomes**[3], **Sara Ferreira**[4]

[1]Federal University of Ceará, Ceará – Brazil, mateus.nogueira@det.ufc.br
[2]Federal University of Ceará, Ceará – Brazil, flaviocunto@det.ufc.br
[3]Federal University of Cariri, Ceará – Brazil, timbo@det.ufc.br
[3]Faculty of Engineering of the University of Porto, Porto – Portugal, sara@fe.up.pt

Submitted:
10 de abril de 2022
Accepted for publication:
2 de dezembro de 2022
Published:
16 de janeiro de 2023
Editor in charge:
José Reynaldo Setti

Keywords:
Pedestrian.
Severity.
Logit Models.
Random Parameters.

Palavras-chave:
Pedestre.
Gravidade.
Modelos Logit.
Parâmetros aleatórios.

DOI:10.14295/transportes.v31i1.2761

ABSTRACT
In Brazil, pedestrians represent the third largest group of crash victims, after motorcyclists and car occupants. Implementing measures to ensure pedestrian safety and prioritization requires an understanding of the risk factors associated with crash injuries. In this study, a random-parameter logit model was estimated to investigate factors influencing the severity of crashes with pedestrians in urban roads in Fortaleza, Brazil. A sample of 2,660 observations of crashes with pedestrians in the city from 2017 to 2019 was used. The injury severity levels adopted by the Crash Information System (SIAT) were grouped into three categories: mild/moderate, severe and fatal. From the investigated factors, only the variable related to the pedestrian's age over 60 years old obtained a significant random parameter. In this case, the heterogeneity in the observations may be associated, among other factors, to the body's physical fragility and the cognitive function that may differ among individuals in this group. The results showed that the driver's gender and age, the crash site, the motorcycle use, and the presence of speed cameras did not have a significant impact on the severity of crashes with pedestrians. On the other hand, crashes occurring at night, with heavy vehicles, on weekends, and located on roads with higher traffic classification are associated with more severe injuries. The incorporation of unobserved heterogeneity in the estimation of the model's parameters stands out as one of the main contributions of this work.

RESUMO
No Brasil, os pedestres representam o terceiro maior grupo de vítimas de acidentes, depois de motociclistas e ocupantes de automóveis. O emprego de medidas para garantir a segurança de pedestres requer uma compreensão dos fatores de risco associados a lesões em acidentes. Um modelo logit de parâmetros aleatórios foi estimado para investigar fatores que influenciam na severidade dos acidentes com pedestres em vias urbanas de Fortaleza – Brasil. Para isso uma amostra 2660 observações de atropelamentos foi utilizada. Dos fatores investigados apenas a variável referente a idade superior a 60 anos do pedestre obteve um parâmetro aleatório significativo. Nesse caso, a heterogeneidade nas observações pode estar associada, entre outros fatores, a fragilidade física do corpo e a função cognitiva que pode diferir entre os indivíduos desse grupo. Os resultados mostraram ainda que os acidentes ocorridos a noite, com veículos pesados, nos finais de semana e localizadas em vias de maior classificação de tráfego estão associadas a lesões mais graves. A incorporação da heterogeneidade não observada na estimação dos parâmetros do modelo destaca-se como uma das principais contribuições deste trabalho.

1. INTRODUCTION

In Brazil, pedestrians represent the third largest group of crash victims, after motorcyclists and car occupants. The fact that they have a significantly smaller mass compared to motor vehicles and are fully exposed without any protective barrier, makes the chances of suffering more severe injuries, when involved in a crash, to be greater than in the vehicle's occupants (Shinar, 2017). In 2018, according to data from the Unified Health System (DATASUS), 33,000 deaths caused by crashes were recorded. Pedestrians accounted for about 20% of these deaths. This proportion is higher if one considers crashes occurring in urban areas. In the city of Fortaleza, for example, about 40% of all traffic deaths are pedestrians (AMC, 2018).

Efforts to prioritize these users need to be coordinated with strategies that increase their safety. This, in turn, requires an understanding of the risk factors that may be associated with the outcome of injury crashes, in order to allow the identification of unsafe urban space configurations and the definition of appropriate countermeasures for transport policies. A significant number of studies related to pedestrian safety have been conducted to investigate the influence of factors such as road and vehicle attributes, pedestrian and driver characteristics, environmental conditions, among other aspects, on the outcome of injuries (Lee and Abdel-Aty, 2005; Eluru et al., 2008; Jang et al. 2013; Aziz et al., 2013).

Historically, statistical analysis of crash data has been used as the basis for the implementation of road safety policies with the goal of preventing crashes and reducing injury severity. Discrete modeling approaches are commonly employed for injury severity analysis in crashes (Savolainen et al., 2011). Due to the ordered nature of injury severity levels, models such as ordered logit/probit are commonly employed (Lee and Abdel-Aty, 2005; Jang et al., 2013). Binary models (Ma et al., 2009; Ferreira et al., 2017), ordered discrete models (Lee and Abdel-Aty, 2005; Zahabi et al., 2011; Jang et al., 2013) and unordered multinomial discrete models (Tay et al. 2011; Chen and Fan, 2019) are statistical approaches widely used in severity studies.

However, when developing such categorical models, it is important to consider a possible unobserved heterogeneity between the explanatory variables, which may occur due to the fact that some of the many factors that affect the probability of a crash and the injury outcomes are not available in the database of analyst. Mannering et al. (2016) provide a detailed discussion of this problem in the context of crash data and analysis and present available statistical approaches to deal with this unobserved heterogeneity. Inserting this aspect into the model increases confidence in the overall structure of the model and its explanatory power.

The assessment of risk factors associated with the severity of crashes with pedestrians has been explored in a very incipient way in the Brazilian road context. Recently Torres et al. (2017) analyzed the prevalence of severity-related factors for vulnerable users (pedestrian or cyclist), but restricted to the surroundings of schools, in the city of Porto Alegre. Having an assessment of these factors for the entire urban environment of a city helps researchers and practitioners to develop effective countermeasures to prevent fatalities or reduce the severity of injuries a pedestrian may suffer when involved in a crash. In this sense, this paper aims to identify and evaluate factors that influence the severity of pedestrians in the city of Fortaleza/Brazil, considering possible unobserved heterogeneity among variables.

2. LITERATURE REVIEW

During a run-over, injuries are consequences of the transfer of energy to the human body and its tolerance to impact. This energy is the result of the vehicle's speed at the time of the crash and its mass. The pedestrian's tolerance to impact, on the other hand, is linked to the physical conditions of the human body to resist the transfer of energy and the area of the body affected (Corben et al., 2004; Yang, 2005). These aspects, however, are not measured or collected directly and are likely not available in the crash databases. Thus, road safety studies seek to identify risk factors that may reflect the effects of these unmeasured aspects on the outcome of injuries. These factors include characteristics of pedestrians and drivers, attributes of the road network, aspects related to the moment of the crash, among others.

Of the pedestrian characteristics, the age of these users is always one of the most notable factors. Researchers have observed that older pedestrians (over 65 years of age at least) and very young pedestrians such as children are associated with an increased risk of serious or fatal injuries when involved in a crash (Aziz et al., 2013; Chen and Fan, 2019; Eluru et al., 2008; Jang et al., 2013).

Some studies have focused on aspects of urban space such as the effect of type of traffic control on crash severity levels and have found that injury levels increase in the absence of traffic controls such as pedestrian traffic lights (Lee and Abdel-Aty, 2005; Eluru et al., 2008). Other specific attributes such as the number of lanes, presence of a central median, land use close to the crash site are frequently reported in the literature. In addition, aspects related to the time of the crash such as lighting conditions, time of day and day of the week were shown to influence the severity of injuries in previous studies (Chen and Fan, 2019; Aziz et al., 2013; Tay et al., 2011; Pour-Rouholamin and Zhou, 2016).

Even when speed is not the decisive cause for a crash to occur, it is highly related to the severity of injuries. Despite its importance, this variable is extremely difficult to observe at the time of occurrence and therefore proxy variables such as road speed or road classification are commonly used (Jang et al., 2013; Li et al., 2016). Previous studies have examined the impact of road speed limits and vehicle type and found that roads that allow higher speeds as well as vehicles of greater mass and dimension are related to greater severity of injuries to pedestrians (Chen and Fan, 2019; Aziz et al., 2013; Eluru et al., 2008; Sze and Wong, 2007).

The influence of these factors on the severity of road traffic injuries in general is analyzed using categorical models. Due to the ordered nature of injury severity, models such as ordered logit/probit are commonly employed (Lee and Abdel-Aty, 2005; Jang et al., 2013). One of the assumptions of this type of approach is that the parameter estimates are constant across the different severity levels. This condition is known as proportional odds or parallel regression assumption (Savolainen et al., 2011). If violated, a viable alternative is the generalized ordered logit model that relaxes this assumption and generalizes the traditional ordered logit model (Cunto and Ferreira, 2017). Another alternative is the use of unordered models, which do not consider the ordinal nature of the response variable but relax the assumption that the impact (for value and sign of the parameters) of the variables is the same at all injury severity levels.

Another assumption often violated with traditional discrete models (both ordered and unordered) is that the parameter estimates obtained are considered as applicable to all observations without considering the influence of unobserved factors. The literature treats this as an unobserved heterogeneity problem. If unobserved heterogeneity is ignored, the model may be incorrectly specified, and the estimated parameters will be generally biased and inefficient, potentially leading to inconsistent inferences (Kim et al., 2010; Mannering et al., 2016). Previous studies have addressed unobserved heterogeneity by allowing parameter estimates to vary randomly across observations. This model is commonly defined as a multinomial random parameter model or mixed logit model, and this parameter variation depends on a distribution defined by the analyst.

Islam and Jones (2014) investigated factors influencing pedestrian injury severity in urban and rural areas in Alabama – USA, by incorporating the effects of randomness into the observations. The estimation results showed that the parameters could be modeled as random parameters, indicating their variable influences on injury severity. The consideration of such randomness has special significance for pedestrian safety because the effect of randomness is likely more pronounced for vulnerable road users such as pedestrians (Kim et al., 2008; Kim et al., 2010). In addition, the authors observed that there are clear differences in the influence of the variables in urban and rural areas. Some variables were significant only in one area (rural or urban), but not in the other.

Haleem et al. (2015) also used the mixed logit model approach to investigate the contributing factors affecting the pedestrian injury severity at signalized and non-signalized intersections. The study used three years of crashes data with pedestrians in Florida (USA). The results showed that at signalized intersections, percentage of trucks, speed limit, and very young pedestrians were found to be statistically significant random parameters. Whereas at non-signalized intersections, old pedestrians and crashes occurring in dark lighting conditions with streetlights were significant random parameters.

Table 1 – Studies that considered unobserved heterogeneity in the analysis of the severity of crashes with pedestrians

Authors	Approach Employed	Investigated Factors
Chen and Fan (2019)	Mixed logit	Driver characteristics (State, age, gender, alcoholic status, physical condition); vehicle type; pedestrian age; road surface; road characteristics (local, roadway class, pavement type, terrain); speed limit; AADT; time of day; day of week; weather condition; rural/urban area; light condition; work zone.
Haleem et al. (2015)	Mixed logit	Speed limit; percentage of trucks; AADT; time of day; weather conditions; presence of pedestrian signals; hour of crash; road surface condition; land use; presence of pedestrian refuge area; crosswalk type; pedestrian age; pedestrian maneuver before crash; driver’s vehicle type
Islam, S. and Jones, S. L. (2014)	Random parameter logit	Pedestrian age and gender; time of day; weather conditions; land use; pedestrian behavior during the crossing; road characteristics; type of control and lighting.
Kim et al. (2010)	Mixed logit/ Random parameter logit	Pedestrian characteristics (age and gender); driver characteristics (age, gender, has been drinking); time of day; traffic control type; land use; time of day; weather conditions; vehicle type; roadway characteristics (road class, geometry, road type); speeding-involved crash; pedestrian behavior; driver maneuver.

Chen and Fan (2019) developed mixed logit models to investigate and identify significant contributing factors to pedestrian injury severity in rural and urban areas in North Carolina (USA). The results indicated that factors such as poor physical condition of the driver, heavy trucks, low lighting conditions, speed limit between 60 and 80 km/h, and speed limit above 80 km/h significantly increase pedestrian injury severity in both rural and urban areas. In rural areas, driver age, pedestrians aged between 26 and 65 years, terrain and speed limit were considered significant random parameters. In urban areas, pedestrian age, road class and speed limit were significant random parameters. Table 1 shows researches that considers the unobserved heterogeneity in the study of the severity of crashes with pedestrians and the risk factors investigated.

For Wang et al. (2013), even in developed countries or cities, differences in road infrastructure, traffic conditions, pedestrian and driver behavioral patterns can result in a different set of significant factors associated with the severity of injuries sustained by pedestrians. Furthermore, the choice of the most appropriate type of model, as well as the selection of factors related to the severity of injury in pedestrians, depends very much on the circumstances of the study site, the dataset used for the analysis and the research objectives. In this study, data from the Brazilian city of Fortaleza will be used to investigate the risk factors associated with the severity of injuries in pedestrians due to the possibility of accessing the data necessary to carry out this study.

3. CRASH DATA AND URBAN ROAD ATTRIBUTES

Crash data with pedestrians were collected from the Crash Information System – Fortaleza (SIAT / FOR) for the years 2017 to 2019. This data consists of individual records with personal information of the victims, characteristics of the vehicles involved, and aspects related to the crash. To determine the variables related to the urban structure, a georeferenced database of traffic lights and speed enforcement cameras provided by the municipality's traffic management agency (AMC) was used. In the crashes dataset, observations with fields without information related to the investigated explanatory variables were eliminated from the initial sample (N = 4,658), resulting in a sample of 2,660 observations. Regarding the severity of injuries suffered by the pedestrian when involved in the crash, SIAT-FOR uses five classification levels: unharmed (no apparent injury), mild (possible injury), moderate (evident injury), severe (disabling injury), and fatal (killed). The uninjured category was eliminated because it presented only one record for the period. The mild and moderate injuries were grouped into a single category due to the similarity in terms of the crash consequence, which can lead to differences in the classification between these two levels by the responsible agents. Thus, the dependent variable was divided into three severity categories: mild/moderate, severe, and fatal.

Table 2 – Description of Variables

Variable	Description
Severity	0 - Mild/Moderate (80%); 1 - Severe (12%); 2 - Fatal (8%)
Pedestrian	
Gender_male	1 – Male (66%); 0 – Female (34%)
Age_0_15	1- Pedestrian up to 15 years old (7%); 0 – Other
Age_16_30	* Pedestrian between 16 and 30 years old (21%);
Age_31_60	1 - Pedestrian between 31 and 60 years old (52%); 0 – Other
Age_M60	1 – Pedestrian above 60 years old (20%); 0 – Other
Driver	
Gender_D_male	1 – Male (84%); 0 – Female (16%);
Age_D18_30	* Driver between 18 and 30 years old (45%);
Age_D31_60	1 - Driver between 31 and 60 years old (52%); 0 – Other;
Age_D60	1- Driver above 60 years old (3%); 0 – Other.
Crash	
Weekend	1 - Weekend (30%); 0 – Working day (70%);
Time_night	1 – From 6 pm to 5 am (43%); 0 – From 5 am to 6 pm (57%);
Mid_block	1 – Middle of the block (83%); 0 – Intersection (17%);
Auto	*Passenger vehicle (46%);
Heavy_Veh	1 – Heavy vehicle (8%); 0 – Other;
Moto	1 – Motorcycle (46%); 0 – Other;

Table 2 – Description of Variables (continuation)

Variable	Description
Road Network Attributes	
R_Loc	*Local road (49%);
R_Col	1 – Collector road (4%); 0 - Other;
R_Art	1 - Arterial road (35%); 0 – Other;
R_Exp	1 - Expressway (12%); 0 – Other;
P_Speed_Cam	1 – Presence of speed cameras in a radius of 100 meters from the crash site (6%); 0 – Absence;
P_Traf_Lights	1 – Presence of traffic lights in a radius of 100 meters from the crash site (30%); 0 – Absence;

*Reference Category

Some variables were obtained from the georeferenced base. For this, a buffer with a radius of 100 meters around the crash was created. Previous studies have worked with buffers between 50 and 600 meters around the crash to collect variables related especially to the built environment or land use (Miranda-Moreno et al., 2011; Zahabi et al., 2011; Prato et al., 2017). In order to verify the presence of traffic lights or speed cameras and knowing that generally drivers reduce speed only when approaching the equipment and then increase speed again, it was decided to consider in this study a buffer of 100 meters radius which could represent the area of influence of these equipment. The independent variables were divided into four groups: pedestrian characteristics, driver characteristics, crash conditions, and road network attributes. Table 2 presents the selected variables, along with their relative frequencies.

4. MODEL DESCRIPTION

As stated previously, the random parameters logit model (RPM) can be developed to incorporate unobserved heterogeneity in its estimates. Just like multinomial logit model (MNL), RPM also defines the utility function of the injury level *i* as:

$$S_{in} = \beta_i X_{in} + \varepsilon_{in} \quad (1)$$

where β_i is a vector of estimated parameters, X_{in} is a vector of observable characteristics that affect the severity of the injury suffered by observation n, and ε_{in} is an error term that accounts for unobserved effects and is assumed to be identically and independently distributed (Washington et al., 2003). As in MNL, when an RPM is estimated, an injury level is used as the comparison group and therefore its coefficients are set to zero. In this study, the first category (mild/moderate) was used as the reference category.

To arrive at the mixed logit model, random parameters are introduced with a density function $f(\beta|\varphi)$, where φ is a vector of parameters describing the density function. This density function can take on a wide range of distributions. The most commonly employed are the uniform, triangular, normal, and log-normal distributions. For this work, the best statistical fit was achieved using the normal distribution. Based on Equation 1, the probability equation for injury severity can be written as:

$$P_{in}(i) = \int \frac{e^{\beta_i X_{in}}}{\sum_I e^{\beta_i X_{in}}} f(\beta|\varphi)\, d\beta \quad (2)$$

The probabilities estimated by RPM are a weighted average for different values of β_i across observations, where some elements of the vector β_i may be fixed and some may be randomly distributed. The estimation of the random parameter logit model shown in Equation 2 can be performed using simulated maximum likelihood approaches, where the probabilities are approximated by drawing values of β_i from $f(\beta|\varphi)$ for given values of φ.

Previous studies have shown that using Halton's sequence produces approximations for numerical integrations that are more accurate than purely random drawings (Train, 2013;

Gkritza and Mannering 2008). For this work, simulations with 100, 150, and 200 Halton draws were tested. The results showed a consistency in significant parameters across simulations. Thus, we chose to estimate the models with simulations of 100 Halton draws as they require less computational effort.

Following the work of Kim et al. (2013), each variable was tested in the model as fixed or random. This process was carried out gradually, starting from the traditional multinomial logit model with all fixed variables. The order of insertion of variables with random parameters was based on the statistical significance of each of the variables in the traditional multinomial model. The decision about the randomness of the parameters is based on both their statistical significance and the significance of their dispersion parameter. In this sense, a random parameter is used when the mean and standard deviation of the parameter is statistically significant, otherwise the parameters are fixed in all observations. This criterion was also adopted in the works of Cunto and Ferreira (2017) and Ferreira et al. (2017).

To evaluate the effect of the variables on the probabilities of each severity level, the marginal effects of each predictor can be calculated. One method for exploring these marginal effects in categorical binary variables is to calculate the change in the predicted probability when a variable is switched from 0 to 1, or vice versa. This result is called direct pseudo-elasticity and is expressed by Equation 3 (Washington et al., 2003).

$$E = \frac{P[Y=i|x=1] - P[Y=i|x=0]}{P[Y=i|x=0]} \quad (3)$$

5. RESULTS AND DISCUSSIONS

The main advantage of the random-parameter model is to provide a more accurate interpretation of the effects of variables on pedestrian injury severity, by capturing the heterogeneity of the observations. The estimates for the random-parameter model are presented in Table 3. Meanwhile, Table 4 provides the pseudoelasticity for each of the variables. The variables that were not significant were removed from the model and therefore do not appear in Table 3. In this sense, the variables related to the drivers' characteristics (sex and age of the driver), crash location, crashes involving motorcycles and the presence of speed cameras do not form statistically significant in the model. It is difficult to point to a specific reason for obtaining a non-statistically significant variable – perhaps because of the data, that is, the number of observations in a category may be too small to be able to estimate it, as the variable referring to the presence of radars of velocity, or simply because the variable may not really affect the dependent variable, which in this case is divided into three categories.

Regarding the characteristics of pedestrians, the variable *Age_M60*, referring to the group of pedestrians over 60 years of age, obtained significant random parameters. This means that the model estimated a value for each observation and these values follow an assumed normal distribution with mean and standard deviation (SD), which are presented in Table 3: 1.808 and 1.653, respectively. The heterogeneity in the observations may be associated in this case with, among other factors, the physical fragility of the body and the cognitive function that may differ among individuals in this group. A 61-year-old pedestrian could have a higher resistance to impact or a shorter reaction time than an 85-year-old pedestrian, for example. The results in Tables 3 and 4 further indicate that pedestrians between the ages of 31 and 60 increase the likelihood of fatal injuries in a crash compared to younger pedestrians. It is not easy to point out direct reasons for this result, which may be associated with a reduction in the physical resilience of victims with advancing age, since this group covers a range from 31 to 60 years of age, with a population with diverse features. On the other hand, this result may also be related to a greater exposure of this group to crashes, considering that this group corresponds to approximately 50% of the victims of crashes in the city, according to data from the Annual Report on Road Safety (AMC, 2018).

The variable *Gender_male* was significant only for the fatal category, with a positive coefficient, meaning that male pedestrians are more likely to suffer fatal injuries when involved in a crash. This increase, according to Table 4 is about 88% more likely than for females. One hypothesis for this result may be related to a tendency of riskier behavior on the part of men, especially in relation to the gaps available for crossing and when the speed of the vehicles is higher, leading to high energy impacts. In the case of Fortaleza, Torres et al. (2020) observed that men tend to perform more aggressive/risky crossings than women at mid-block crossings.

Table 3 – Results of the Estimate of the Random Parameter Model

	Severe		***Fatal***	
	coef	*p value*	*coef*	*p value*
Intercept	-2.802	0.000*	-5.494	0.000*
Gender_male	0.059	0.664	0.636	0.002*
Age_31_60	0.226	0.125	0.942	0.000*
Age_M60	0.063	0.754	1.808	0.001*
sd(Age_M60)			1.653	0.062***
Gender_D_male	-0.053	0.752	0.574	0.034**
Weekend	0.326	0.015**	0.447	0.015*
Time_night	0.507	0.000*	0.595	0.001*
Heavy_Veh	0.949	0.000*	1.496	0.000*
R_Exp	1.491	0.000*	1.700	0.000*
R_Art	0.426	0.006*	0.722	0.001*
R_Col	0.449	0.020**	0.327	0.241
P_Traf_Lights	-0.300	0.046**	-0.061	0.747
Log Likelihood at convergence	-1,540.4			
Log Likelihood at zero	-2,922.3			
Pseudo R^2	0.47			

*Statistically significant at α = 0.01; **statistically significant at α = 0.05;
***statistically significant at α = 0.1;

Table 4 – Pseudoelasticity

	Mild/Moderate	***Severe***	***Fatal***
Gender_male	0%	-	88%
Age_31_60	-1%	-	155%
Age_M60	-2%	-	498%
Gender_D_male	0%	-	77%
Weekend	-2%	35%	53%
Time_night	-4%	59%	74%
Heavy_Veh	-9%	134%	304%
R_Exp	-18%	266%	351%
R_Art	-3%	48%	99%
R_Col	-3%	52%	-
P_Traf_Lights	2%	-25%	-

- non-significant for a confidence level of 0.10

For driver characteristics, only gender was significant. As shown in Table 3, male drivers are more likely to increase the likelihood of pedestrian fatality compared to female drivers. Eluru et al. (2008) and Kim et al. (2008) point out that this result may have a relationship with a higher risk-taking behavior, to assume higher speeds of this group of drivers compared to female drivers. According to data from the Federal Highway Police, in the year 2019, considering the crashes that occurred on the highways that cross the urban perimeter of Fortaleza and whose main cause was incompatible speed, 85% involved male drivers (PRF, 2019).

Regarding the aspects linked to the crash time, the results in Table 3 indicate an increase in the chances of severe and fatal injuries at night. According to the pseudoelasticities, a crash with pedestrians at night has 59% and 74% more chances of resulting in severe and fatal injuries to the pedestrian, respectively. In addition, a crash with pedestrians on a weekend has about 35% more chances of resulting in a severe injury and 53% more chances of resulting in a fatal injury. Similar results were observed in the work of Chen and Fan (2019) who used mixed models and in the works of Jang et al. (2013), Zafri et al. (2020), and Batouli et al. (2020) who adopted logistic regression models. High energy impact events for the pedestrian may be a result, at least in part, of inappropriate speed choice made by drivers under the influence of substances, possibly on weekends, as well as delayed responses caused by their impaired condition or due to lower visibility conditions in the evening.

As for the attributes of the road network, the functional road class is commonly used as a proxy variable for the average speed of the involved vehicles. Although the Brazilian Traffic Code establishes a limit of 80 km/h for express roads, in Fortaleza the limit of 60 km/h is adopted for these roads as well as for arterial roads in which the pedestrian crossing is allowed. According to Table 3, both arterials and express roads are associated with a higher risk of severe and fatal pedestrian injuries, which is notably attributed to the higher speeds achieved by the drivers. The chances of fatal injuries increase by 99% on arterials and 351% on expressways compared to local roads, according to Table 4. In the context of Fortaleza, these findings may be associated with a considerably lower traffic flow at some periods of the day or week, which leads to higher chances of events where the vehicle speed was too high for the pedestrian's body, even if (some of them) were below or close to the speed limit. In addition, it is important to note that on these roads, the presence of heavy vehicles, especially freight vehicles, is also more common. The combination of these two factors can further increase the risk of more severe crash outcomes.

The variable *Heavy_Veh* which include trucks, vans, and buses, significantly influences the severity of injuries, increasing the chances of more severe injuries to the pedestrian. Besides vehicle mass, which has a decisive effect on the energy transferred during impact, the shape of these larger vehicles can also be related to the concentration of energy that is transferred to the pedestrian upon the impact the affected area of the victim's body. This relationship was found in previous studies such as those by Aziz et al. (2013), who used random parameter models, and Wang et al. (2013) who used logistic regression models.

The negative sign of the coefficient of the variable *P_Traf_Lights* points out that the presence of traffic lights near the crash site reduces the chances of severe injuries to pedestrians. Aziz et al. (2013), who used random parameter models, and Sze and Wong (2007), who used logistic regression models, obtained similar results in their work. The authors attribute these results to better priority indication at locations with traffic control which leads to greater caution on the part of both drivers and pedestrians.

6. FINAL REMARKS

This study aimed to identify and analyze risk factors associated with the severity of crashes with pedestrians in urban areas in the Brazilian road context, in order to provide evidence to road safety policymakers to address and treat problems related to pedestrian safety.

To this end, random parameter logit models were estimated using a sample with 2,660 observations of crashes with pedestrians collected from the Crash Information System – Fortaleza (SIAT-FOR) for the years 2017 to 2019.

The results of this research identified several risk factors associated with pedestrian and driver characteristics, attributes of the road network, and aspects related to crash conditions. A traditional multinomial logit model was improved by introducing heterogeneity in parameter effects through random parameter logit models considering a normal distribution form and Halton drawings for parameter estimation. The final structure of the model showed individual heterogeneity only in the variable referring to the group of pedestrians over 60 years old. The other variables were treated as fixed.

Regarding human factors, the age and gender of the pedestrian and driver involved were investigated. The estimated model suggested that elderly pedestrians (over 60 years old) are on average 6 times more likely to be fatally injured in crashes compared to young pedestrians (between 16 and 30 years old). The driver's age, on the other hand, did not show statistical significance for injury outcomes. As for the gender of drivers and pedestrians, the chances of fatal injuries to the pedestrian increase when the driver is male and also when the pedestrian himself is male.

Factors related to the crash such as the time of day (day/night) as well as the day of the week had a significant influence on the crash outcome. A trampling at night is associated with higher chances of severe injuries to pedestrians. Weekend crashes, on the other hand, are responsible for a higher probability of fatal injuries. Some operational attributes of the road network were also investigated, such as the presence of speed traps and traffic lights around the crash site. The results indicated that traffic lights reduce the chances of severe injuries to pedestrians.

Heavy vehicles such as buses, trucks, vans, among others, are associated with a higher probability of severe crashes with pedestrians. In addition, higher-rated roads, such as express or arterial roads, characterized by being wider and having more lanes, have been shown to be associated with a higher likelihood of more severe crashes with pedestrians. It is worth noting that in 2018 the city of Fortaleza began a process of adapting the speed limits of some arterial roads from 60 to 50 km/h.

The results of this work have increased the general knowledge about pedestrian-related crashes. It is well known that the mass and speed of the vehicle in the moment when the pedestrian crash are crucial to the severity outcome. The potential influence of speed as a factor was indirectly inferred using time of the day, day of the week, and roadway classification as proxy variables. While it may be incredibly challenging to capture individual vehicle speeds for each occurrence, it may be possible to use a more informative variable such as average traffic flow speed obtained either from historical loop detector data or by micro- and mesoscopic simulation.

An additional layer of analysis using road user attributes that could help to explain the (inappropriate) speed choice, such as risky behavior after alcohol consumption, for example, was not fully possible as this information was not available in the dataset. Incorporating these aspects would likely support better the identification and isolation of the main factors influencing crash severity during the weekend and evening. In addition, it may be useful to investigate how temporal and spatial dimensions influenced the overall results obtained in this study.

As a limitation of this study, the loss of many observations from the database stands out. Information on crashes in Fortaleza comes from different sources and is consolidated by technicians from the municipality's traffic management agency. However, technicians do not always get all the information about the event, which ends up reducing the number of observations with data needed to carry out studies like this one.

ACKNOWLEDGEMENTS
This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) of Brasil under Finance Code 001, the National Council for Scientific and Technological Development (CNPq), and the Ceará Foundation for Scientific and Technological Development (FUNCAP). The authors acknowledge the financial support provided by these agencies.

REFERENCES

AMC (2018) Relatório Anual de Segurança Viária de Fortaleza 2018. Autarquia Municipal de Trânsito e Cidadania, Fortaleza, 2018.

Aziz, H; M. A.; S. V. Ukkusuri and S. Hasan (2013) Exploring the determinants of pedestrian-vehicle crash severity in New York City. *Accident Analysis and Prevention*, v. 50, p. 1298–1309, 2013. doi:10.1016/j.aap.2012.09.034

Batouli, G.; M. Guo; B. Janson and W. Marshall (2020) Analysis of pedestrian-vehicle crash injury severity factors in Colorado 2006–2016. *Accident Analysis and Prevention*, v. 148. doi: 10.1016/j.aap.2020.105782

Chen, Z. and Fan, W. (2019) Modeling Pedestrian Injury Severity in Pedestrian-Vehicle Crashes in Rural and Urban Areas: Mixed Logit Model Approach. *Transportation Research Record*, v. 2673, p. 1023–1034. doi: 10.1177/0361198119842825

Corben, B.; M. Cameron; T. Senserik and G. Rechnitzer (2004). Development of the visionary research model: application to the car/pedestrian conflict. (Rep n. 229). Melbourne: Monash University Accident Research Centre.

Cunto, F. J. and S. Ferreira (2017) An analysis of the injury severity of motorcycle crashes in Brazil using mixed ordered response models. *Journal of Transportation Safety & Security*, 9(sup1), 33-46. doi: 10.1080/19439962.2016.1162891

Dutta, B. and V. Vasudevan,(2017) Study on pedestrian risk exposure at unsignalized intersection in a country with extreme vehicle heterogeneity and poor lane discipline. *Transportation Research Record*, v. 2634, p. 69–77. doi: 10.3141/2634-11

Eluru, N.; C. R. Bhat and D. A. Hensher (2008) A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention*, v. 40, n. 3, p. 1033–1054. doi: 10.1016/j.aap.2007.11.010

Ferreira, S.; M. Amorim and A. Couto (2017) Risk factors affecting injury severity determined by the MAIS score. *Traffic Injury Prevention*, v. 18, p. 1–29. doi: 10.1080/15389588.2016.1246724

Gkritza, K. and F. L. Mannering (2008) Mixed logit analysis of safety-belt use in single- and multi-occupant vehicles. *Accident Analysis and Prevention*, v. 40, p. 443–451. doi: 10.1016/j.aap.2007.07.013

Haleem, K.; P. Alluri and A. Gan (2015) Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis and Prevention*, v. 81, p. 14–23. doi: 10.1016/j.aap.2015.04.025

Islam, S. and S. L. Jones (2014) Pedestrian at-fault crashes on rural and urban roadways in Alabama. *Accident Analysis and Prevention*, v. 72, p. 267–276. doi: 10.1016/j.aap.2014.07.003

Jang, K.; S. Park; S. Kang; K. K. Song and S. Chung (2013) Evaluation of pedestrian safety. *Transportation Research Record*, n. 2393, p. 104–116. doi: 10.3141/2393-12

Kim, J. K.; G. F. Ulfarsson; V. N. Shankar and S. Kim (2008) Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic logit analysis. *Accident Analysis and Prevention*, v. 40, n. 5, p. 1695–1702. doi: 10.1016/j.aap.2008.06.005

Kim, J. K.; G. F. Ulfarsson; V. N. Shankar and F. L. Mannering (2010) A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis and Prevention*, v. 42, n. 6, p. 1751–1758. doi: 10.1016/j.aap.2010.04.016

Lee, C. and M. Abdel-Aty (2005) Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accident Analysis and Prevention*, v. 37, n. 4, p. 775–786. doi: 10.1016/j.aap.2005.03.019

Li, D; P. Ranjitkar; Y. Zhao; H. Yi and S. Rashidi (2016) Analyzing pedestrian crash injury severity under different weather conditions. *Traffic Injury Prevention*,v.18, n.4, p. 427–430. doi: 10.1080/15389588.2016.1207762

Ma, Z.; C. Shao; H. Yue and S. Ma (2009) Analysis of the logistic model for accident severity on urban road environment. *Intelligent Vehicles Symposium*, Proceedings, p. 983–987. doi: 10.1109/IVS.2009.5164414

Mannering, F. L.; V. Shankar and C. R. Bhat (2016) Unobserved heterogeneity and the statistical analysis of highway accident data. *Analytic Methods in Accident Research*, v. 11, p. 1–16. doi: 10.1016/j.amar.2016.04.001

Miranda-Moreno, L. F.; P. Morency and A. M. El-Geneidy (2011) The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, v. 43, n. 5, p. 1624–1634. doi: 10.1016/j.aap.2011.02.005

Pour-Rouholamin, M. and H. Zhou (2016) Investigating the risk factors associated with pedestrian injury severity in Illinois. *Journal of Safety Research*, v. 57, p. 9–17. doi: 10.1016/j.jsr.2016.03.004

Polícia Rodoviária Federal (2019): banco de dados. Available at: https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos/dados-abertos-acidentes. (visited on 5/mar/2022)

Prato, C. G.; S. Kaplan; A. Patrier and T. K. Rasmussen (2018) Considering built environment and spatial correlation in modeling pedestrian injury severity. *Traffic Injury Prevention*, v. 19, n. 1, p. 88–93. doi: 10.1080/15389588.2017.1329535

Rosenbloom T. (2009) Crossing at a red light: Behaviour of individuals and groups. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 12, n. 5, p. 389–394. doi: 10.1016/j.trf.2009.05.002

Savolainen, P. T.; F. L. Mannering; D. Lord and M. A. Quddus (2011) The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention*, v. 43, n. 5, p. 1666–1676. doi: 10.1016/j.aap.2011.03.025

Shinar, D. (2017). *Traffic safety and human behavior* (2nd ed). Emerald Group Publishing.

Sze, N. N. and S. C. Wong (2007) Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes. *Accident Analysis and Prevention*, v. 39, n. 6, p. 1267–1278. doi: 10.1016/j.aap.2007.03.017

Tay, R.; J. Choi; L. Kattan and A. Khan, (2011) A multinomial logit model of pedestrian–vehicle crash severity. *International Journal of Sustainable Transportation*, v. 5, n. 4, p. 233-249. Doi: 10.1016/j.ijtst.2018.10.001

Torres, C.; L. Sobreira; M. Castro-Neto; F. Cunto; A. Vecino-Ortiz; K. Allen; A. Hyder and A. Bachani (2020) Evaluation of pedestrian behavior on mid-block crosswalks: a case study in Fortaleza- Brazil. *Frontiers in Sustainable Cities*, v. 2, p. 1–6. doi: 10.3389/frsc.2020.00003

Torres, T. B.; A. M. L. Uriarte; C. P. Demore and C. T. Nodari (2017) Prevalência de fatores associados à severidade dos acidentes em entorno de escolas. *Transportes*, v. 25, n. 3, p. 102. doi: 10.14295/transportes.v25i3.1331

Train, K. E. 2003. *Discrete choice methods with simulation*, (2nd ed.). Cambridge University Press.

Wang, Y. Y.; M. M. Haque; H. C. Chin and J. G. J. Yun (2013) Injury severity of pedestrian crashes in Singapore. *In Australasian Transport Research Forum*, ATRF 2013 – Proceedings.

Washington, P. S.; G. M. Karlaftis and F. L. Mannering (2003) *Statistical and Econometric Methods for Transportation Data Analysis*. Chapman & Hall/CRC, Nova Iorque, 2003.

Yang, J. (2005) Review of injury biomechanics in car-pedestrian collisions. *International Journal of Vehicle Safety*, v. 1, n. 1–3, p. 100–117. doi: 10.1504/IJVS.2005.007540

Ye, F. and D. Lord (2014) Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. *Analytic Methods in Accident Research*, v. 1, p. 72–85. doi: 10.1016/j.amar.2013.03.001

Zafri, N. M.; A. A. Prithul; I. Baral and M. Rahman (2020) Exploring the factors influencing pedestrian-vehicle crash severity in Dhaka, Bangladesh. *International Journal of Injury Control and Safety Promotion*, v. 27, n. 3, p. 300-307. doi: 10.1080/17457300.2020.1774618

Zahabi, S. A. H.; K. Manaugh and L.F. Miranda-Moreno (2011) Estimating potential effect of speed limits, built environment, and other factors on severity of pedestrian and cyclist injuries in crashes. *Transportation Research Record*, n. 2247, p. 81–90. doi: 10.3141/2247-10