A home-based approach to understanding seatbelt use in single-occupant vehicles in Tennessee
Application of a latent class binary logit model

Afghari, Amir Pooyan; Hezaveh, Amin Mohamadi; Haque, Md Mazharul; Cherry, Christopher

DOI
10.1016/j.aap.2020.105743

Publication date
2020

Document Version
Final published version

Published in
Accident Analysis and Prevention

Accident Analysis and Prevention 146 (2020) 105743

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

A home-based approach to understanding seatbelt use in single-occupant vehicles in Tennessee: Application of a latent class binary logit model

Amir Pooyan Afghari[a,*], Amin Mohamadi Hezaveh[b], Md Mazharul Haque[c], Christopher Cherry[c]

[a] Safety and Security Science Section, Faculty of Technology, Policy and Management, Delft University of Technology, Netherlands
[b] Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, United States
[c] School of Civil and Environmental Engineering, Science and Engineering Faculty, Queensland University of Technology, Brisbane, QLD, 4001, Australia

ARTICLE INFO

Editor: Chennai Guest Editor

Keywords:
Seatbelt use
Location-based characteristics
Latent class model
Transport safety
Driving behaviour

ABSTRACT

Although the enforcement of seatbelt use is considered to be an effective strategy in reducing road injuries and fatalities, lack of seatbelt use still accounts for a substantial proportion of fatal crashes in Tennessee, United States. This problem has raised the need to better understand factors influencing seatbelt use. These factors may arise from spatial/temporal characteristics of a driving location, type of vehicle, demographic and socioeconomic attributes of the vehicle occupants, driver behaviours, attitudes, and social norms. However, the above factors may not have the same effects on seatbelt use across different individuals. In addition, the behavioural factors are usually difficult to measure and may not always be readily available. Meanwhile, residential locations of vehicle occupants have been shown to be associated with their behavioural patterns and thus may serve as a proxy for behavioural factors. However, the suitability of geographic and residential locations of vehicle occupants to understand the seatbelt use behaviour is not known to date.

This study aims to fill the above gaps by incorporating the residential location characteristics of vehicle occupants in addition to their demographics and crash characteristics into their seatbelt use while accounting for the varying effects of these factors on individual seatbelt use choices. To achieve this goal, empirical data are collected for vehicular crashes in Tennessee, United States, and the home addresses of vehicle occupants at the time of the crash are geocoded and linked with the census tract information. The resulting data is then used as explanatory variables in a latent class binary logit model to investigate the determinants of vehicle occupants' seatbelt use at the time of the crash. The latent class specification is employed to capture the unobserved heterogeneity in data. Results show that Tennessean drivers belong to two general categories—conformist and eccentric—with gender, vehicle type, and income per capita determining the likelihood of these categories. Overall, male drivers, younger drivers, and drivers who have consumed drugs are less likely to wear a seatbelt, whereas drivers who come from areas with higher population density, travel time, and income per capita are more likely to wear a seatbelt. In addition, driving during the day and in rainy weather are associated with an increased likelihood of seatbelt use. The findings of this study will help developing effective policies to increase seatbelt use rate and improve safety.

1. Background

The effects of wearing a seatbelt in reducing roadway injuries and fatalities have long been documented in the road safety literature (Knapper et al., 1976; Hodson-Walker, 1970). Several studies have shown that proper use of seatbelt increases the likelihood of vehicle occupants' survival from a potentially fatal crash by 44 %–73 %, depending on the seating position and vehicle type (Blincoe et al., 2015). Moreover, seatbelt use enforcement is universally considered to be one of the most effective precautionary measures in reducing road fatalities. Previous research has consistently shown that enforcement can significantly increase seatbelt use among vehicle occupants all over the world (Dee, 1998; Eby et al., 2000). In the United States as well, many studies and reports have documented the effectiveness of enforcement programs

* Corresponding author.
E-mail addresses: a.p.afghari-1@tudelft.nl (A.P. Afghari), amohamad@vols.utk.edu (A.M. Hezaveh), m1.haque@qut.edu.au (M.M. Haque), Cherry@utk.edu (C. Cherry).

https://doi.org/10.1016/j.aap.2020.105743
Received 8 July 2020; Received in revised form 11 August 2020; Accepted 18 August 2020
Available online 28 August 2020

such as ‘Click It or Ticket’ campaigns and saturation patrols (Reinfurt, 2004; Solomon et al., 2004; Thomas et al., 2011, 2008; Tison and Williams, 2010, NHTSA, 2009). However, non-compliance with seatbelt use enforcement still accounts for a considerable proportion of roadway crashes in the United States (Shakya et al., 2020), and the magnitude of this problem is not the same across different states in this country (Morgan, 2015; Thomas et al., 2017). According to the recent report by the National Highway Traffic Administration, the average seat belt use in the state of Tennessee is lower than the national average seatbelt use (NHTSA, 2017). Roadside observations also show that a significant proportion of Tennessean vehicle occupants do not use seatbelt despite its proven effectiveness. Previous studies showed that a substantial portion of Tennessean who died in traffic crashes failed to wear their seat belt properly at the time of the crash (Cherry et al., 2018; Hezaveh and Cherry, 2019a; Hezaveh et al., 2019b). These statistics have raised the need to better understand factors influencing seatbelt use that could, in turn, lead to the deployment of effective countermeasures to encourage seatbelt use among vehicle occupants in general, and in Tennessee in particular.

Past research has shown that seatbelt use is associated with a multitude of factors including spatial and temporal characteristics of driving, type of vehicle, demographic and socioeconomic attributes of the vehicle occupants and their behaviours, attitudes and social norms (Lund, 1986; Ali et al., 2011; Okamura et al., 2012; Hezaveh and Cherry, 2019b). Glassbrenner and Ye (2007), Nichols et al. (2009); Reagan et al. (2013) reported a higher seatbelt use compliance rate in urban areas and expressways in comparison with rural roads. Gkritza and Mannering (2008a) found that drivers of single-occupant vehicles are less likely to use a seatbelt in the morning, while drivers of multi-occupant vehicles are less likely to be restrained in the afternoon. Moreover, several studies have shown that seatbelt use rate is significantly lower during night-time compared to daytime (Chaudhary et al., 2005; Chaudhary and Preusser, 2006; Solomon et al., 2007; Tison et al., 2010; Vivoda et al., 2007). The type of vehicle (e.g., small car, large trucks, buses, caravans, and vans) can also influence seatbelt use. Several studies have reported that occupants of pickup trucks have the lowest seat belt use rate compared to occupants of other types of vehicles (Boyle et al., 2003, Glassbrenner and Ye, 2007, Gkritza and Mannering, 2008b). The latest roadside observation reports in Tennessee (CTR, 2018) are also consistent with this latter finding.

Demographic factors also influence vehicle occupants’ seatbelt use. Numerous studies have shown that males are more prone to not wearing seatbelt compared to females (e.g., Pickrell and Ye, 2009; Gkritza and Mannering, 2008; Hezaveh and Cherry, 2019a). Younger drivers have been reported to have lower seatbelt use rates compared to older drivers (Glassbrenner et al., 2004; Calisir and Lehto, 2002). Drivers with higher education levels and higher income levels have been reported to have higher seatbelt use rates (Houston and Richardson, 2005; Wells et al., 2002). These findings are all aligned with many other self-reported studies of seatbelt use which have shown that gender, education, and income are significantly associated with aberrant driving behaviour (Hezaveh et al., 2017, 2018; Martinussen et al., 2013; Nordfjærn et al., 2015). Furthermore, in the US, roadside observations showed that African-American vehicle occupants have lower seatbelt use rates compared to White or Hispanic vehicle occupants (Pickrell and Ye, 2009; Vivoda et al., 2004; Gkritza and Mannering, 2008b).

The above studies have provided an understanding of how spatial, temporal, and vehicle characteristics, as well as demographic attributes (e.g., gender, age, education level, income) influence vehicle occupants seat belt use. However, a complete list of these attributes may not be readily available, especially from crash reports. In addition, a few studies have suggested that sociodemographic attributes of the vehicle occupants as well as their behaviours, attitudes, descriptive and social norms could also influence their seatbelt use (Lund, 1986; Ali et al., 2011; Okamura et al., 2012; Hezaveh and Cherry, 2019b). However, these psychological attributes are very difficult to measure, and this information is usually not readily available (Afghari et al., 2018). This problem is even more acute, noting that the differences in the psychological and behavioural attributes of individuals may result in the varying effects of other factors on their seatbelt use (Eluru and Bhat, 2007; Mannering et al., 2016). For example, while many studies have shown that male drivers are less likely to wear seatbelt compared to female drivers (e.g., Pickrell and Ye, 2009; Gkritza and Mannering, 2008; Hezaveh and Cherry, 2019a), other studies have shown that some male drivers may be more safety-conscious than female drivers and thus may be more likely to wear a seatbelt (Abay et al., 2013). This heterogeneity in the effect of gender on seatbelt use arises from unobserved behavioural factors (e.g., safety-consciousness) and thus may result in erroneous inferences about the effect of gender if not accounted for in modelling seatbelt use. The unobserved heterogeneity among the factors associated with seatbelt use represents a significant modelling challenge.

In addition to fitting an appropriate econometric model to take into account unobserved heterogeneity, a possible solution to overcome the above challenges could be to incorporate any available information about vehicle occupants’ residential locations into their seatbelt use choices and account for their varying effects on seatbelt use. This information may serve as a proxy for those attributes whose data are not available. Earlier studies have also used similar proxies for driver behaviour (Shaon et al., 2019). In addition, it is intuitive to postulate that vehicle occupants’ seatbelt use is influenced by the geographic location where they come from because geography is associated with behaviours, attitudes, and social norms (Rentfrow, 2010). In fact, the geography of vehicle occupants’ residential location might also serve as a proxy for their behavioural patterns when such data are not available (Van Acker et al., 2010; Foster, 1999; Kamruzzaman and Hine, 2013). However, to the authors’ best knowledge, the effects of geographic and residential location of vehicle occupants on their seatbelt use have been less examined.

This study aims to fill the above research gaps by simultaneously incorporating the residential location characteristics of vehicle occupants in addition to their demographic attributes and crash characteristics into their seatbelt use while accounting for the varying effects of these factors on individual seatbelt use choices. To achieve this goal, empirical data are collected for vehicular crashes in Tennessee, United States, and the home addresses of vehicle occupants at the time of the crash are geocoded and linked with the census tract information. The resulting data is then used as explanatory variables in a latent class binary logit model to investigate the determinants of vehicle occupants’ seatbelt use. The latent class specification captures the varying effects of the explanatory variables on seatbelt use.

We acknowledge that vehicle occupants’ seatbelt use vary by their seating position. The latest roadside observation in Tennessee indicates that drivers have lower seat belt use rate compared to the front row passengers (CTR, 2018); this finding was also supported by police crash reports and phone surveys in Tennessee (Hezaveh et al., 2019a,b; Hezaveh and Cherry, 2019a). However, seatbelt use of vehicle occupants in multi-occupant vehicles may be inter-related. For example, seatbelt use of front and rear passengers may be influenced not only by the seatbelt use of drivers but also by the seatbelt use of each other. In the same fashion, the census tract information of each vehicle occupant may have a shared influence on the seatbelt use of multiple occupants. As such, and for simplicity, this study only focuses on the seatbelt use of vehicle occupants in single-occupant vehicles (i.e., drivers). Investigation of vehicle occupants’ seatbelt use in multi-occupant vehicles with the abovementioned complexities is left for future research. Nonetheless, the average vehicle occupancy is very low in the United States, and thus the majority of the vehicles are single-occupant (Office Of Energy Efficiency and Renewable Energy, 2018).

The remaining of this paper is organized as follows: empirical data processing and geocoding are presented in the next section, followed by the details of the latent class logit modelling methodology. The model results are then presented and discussed. Finally, the conclusions and limitations of the study and the future research directions are presented.

2. Methods

2.1. Empirical data processing and geocoding

The empirical data in this study were provided by Tennessee Integrated Traffic Analysis Network (TITAN), which is a state-wide repository for traffic crashes and surveillance reports completed by Tennessee law enforcement agencies. For the year 2016, the data include 247,536 crashes and information about 725,388 drivers who were involved in these crashes. TITAN provides information regarding seatbelt use by occupants at the time of the crash. For this study, we defined seatbelt 'non-use' as vehicle occupants who did not restrain both lap and shoulder seat belt at the time of a crash. Furthermore, the Bing API was used in this study for geocoding the residential addresses of drivers. Only those addresses with an accuracy level of the premise (e.g., property name, building name), address-level accuracy, or intersection level accuracy were used in the analysis (Hezaveh and Cherry, 2019a,b). For more details about the geocoding process of the home addresses and the accuracy, please see Merlin et al. (2020) and Hezaveh et al. (2019a).

After controlling for the number of occupants in vehicles (i.e., excluding multi-occupant vehicles) and cleaning the data (i.e., removing the incomplete records and error entries), 242,468 observations with a Tennessee address were selected for assignment to the census tract data. Census tract data from the U.S. survey in 2010 were also used for obtaining sociodemographic data elements. Table 1 provides summary statistics of driver attributes and crash characteristics and and Table 2 provides summary statistics of residential location characteristics of drivers obtained from census tracts. To prevent outliers, we only considered the census tracts that had more than 20 observations.

2.2. Latent class binary logit model

Binary logit models have been widely used in the statistical literature as the modelling approach to estimate the effects of exogenous factors (e.g., the abovementioned attributes) on individuals' binary choices (e.g., wearing a seatbelt or not) (Washington et al., 2020; Kadilar, 2016). These models are based on the *random utility* theory, according to which individuals make a choice between two (or multiple) alternatives based on observed and unobserved factors. The term random utility is used for this theory because the variation in individuals' behaviour (for example, wearing or not wearing a seatbelt) is due to the randomness that is not explainable by observed factors. As a result, a utility is defined for an individual's choice behaviour that consists of a deterministic and an error term. The deterministic term indicates the systematic effects of observed factors, whereas the error term indicates the random effects of unobserved factors on individual choices. However, an important limitation of the binary logit model is that model parameters in the utility function are assumed to be fixed across individuals. This assumption ignores the possible heterogeneity among individuals in the sample. In fact, the effects of individual attributes such as sociodemographic factors on seatbelt use may not be homogeneous across drivers. This phenomenon is referred to as *unobserved heterogeneity* in the sample (Hensher et al., 2005) and needs to be accounted for within binary logit models.

Table 1
Summary statistics of driver attributes and crash characteristics used in the study.

Variable	Frequency	Sample share
Seat belt use (0= no, 1= yes)	224,573	0.926
Age:		
less than 25 years old	63,284	0.261
between 25 and 40 years	67,164	0.277
between 40 and 60 years old	72,013	0.297
over 60 years old	40,007	0.165
Alcohol consumption (0= no, 1= yes)	5064	0.021
Distraction (0= not distracted, 1= distracted)	14,894	0.061
Drug consumption (0= no, 1= yes)	2735	0.011
Gender: male	128,105	0.528
Day time	182,502	0.753
Weather – clear	183,828	0.758
Weather – rainy	25,713	0.106
Vehicle body type: large passenger vehicles	51,162	0.211

Table 2
Summary statistics of residential location characteristics used in the study and extracted from census tract data.

	Mean	SD	min	Max
Total population density (1000 person/km2)	0.643	0.785	0.000	33.099
Percentage population under 16 years old	0.237	0.077	0.000	0.713
Percentage population above 60 years old	0.185	0.091	0.000	0.894
Percentage of white race	0.764	0.278	0.000	1.000
Average travel time (hrs)	0.428	0.098	0.000	1.097
High school education percentage	0.478	0.193	0.000	0.970
College education percentage	0.216	0.075	0.000	1.000
Median household income ($100,000)	0.508	0.251	0.000	2.494
Income per capita	0.250	0.116	0.000	1.287
Percentage of vacant houses	0.106	0.091	0.000	1.000

Various modelling techniques have been suggested in the literature to address unobserved heterogeneity in transport applications (Mannering et al., 2016). Random parameters specification is perhaps one of the most common approaches in addressing unobserved heterogeneity (Anastasopoulos and Mannering, 2009; Afghari et al., 2016; Fountas et al., 2018; Huang et al., 2019). The parameters in this approach are assigned with random distributions allowing them to vary across observations. This approach has recently been extended to random parameters specification with heterogeneity in the means (Behnood and Mannering, 2017a,b; Hamed and Al-Eideh, 2020; Sharma et al., 2020) and random parameters specification with heterogeneity in the means and variances (Behnood and Mannering, 2017a,b; Seraneeprakarn et al., 2017; Waseem et al., 2019; Heydari et al., 2018) of parameters. Although these elaborate models are theoretically able to capture additional variance in the parameters resulted from unobserved heterogeneity, they are prone to identification problems in that they require enriched empirical data with enough variation in explanatory variables (Mannering at al., 2016). The latent class modelling technique has been introduced as another promising approach to capture the unobserved heterogeneity in the data (Greene and Hensher, 2003)(Park and Lord, 2009; Abolhassani et al., 2019). In the latent class model specification, it is assumed that there are a finite number of classes over the population, and observations are allowed to belong to those classes with different probabilities. This mechanism accounts for possible unobserved heterogeneity that may exist in the data (Mannering et al., 2016). Furthermore, the latent class approach may be combined with the random parameters approach such that the parameters within each class are allowed to vary across observations in that class (Buddhavarapu et al., 2016; Xiong and Mannering, 2013). It is hypothesized in this study that there are latent (unobserved) groups of individuals that share common behaviour depending on their residential location characteristics. As a result, the latent class approach is employed in this study to model drivers' seatbelt use in single-occupant vehicles. It is worth mentioning that the specification of the random parameters within the latent class approach is computationally burdensome due to the large sample size (242,468 records) of this study. As such, the latent class model specification is specified with fixed parameters. The details of this model are presented in the following.

Let i ($i = 1, 2, 3, \ldots, N$) be an index to represent drivers in single-occupant vehicles. The utility of using seatbelt by a given driver (U_i) is stated as:

$$U_i = \boldsymbol{\beta} \boldsymbol{X}_i + \varepsilon_i \tag{1}$$

where U_i is the utility of seatbelt use, $\boldsymbol{\beta}$ is the vector of estimable parameters (including the intercept), $\boldsymbol{X}_i$ is the vector of explanatory variables for individual i (e.g., sociodemographic factors, time of day). ε_i is the random term describing the random part of the utility. Assuming that ε_i is generalized extreme value distributed (McFadden, 1981), the probability of using seatbelt presented by can be stated as:

$$P(seatbelt\ use) = \frac{e^{\beta X_i}}{1 + e^{\beta X_i}} \tag{2}$$

The likelihood of using seatbelt across all individuals can then be determined by the product of Eq. (2) over the entire observations. Such a model specification is referred to as a binary logit model (Washington et al., 2020). This model is now extended into the latent class specification to account for unobserved heterogeneity in data.

Assuming that there are S number of latent classes over the population, the probability of observations belonging to each distinct class, $P(C_s)$, can be computed using a logit model with the following specifications:

$$P(C_s) = \frac{e^{U_s}}{\sum_{s=1}^{S} e^{\beta X_i}} \text{ and } U_s = \boldsymbol{\Omega}_s Z_s \tag{3}$$

where $\boldsymbol{\Omega}$ is the vector of parameters (including an intercept), and $\boldsymbol{Z}$ is the vector of class-specific covariates. Such covariates determine the probabilities of observations being assigned to each specific class. Within each class, the probability of seatbelt use conditioned to that class can be computed using the Eq. (2). Applying the rules of conditional probabilities, the marginal probability of the latent class logit model is stated as:

$$P(seatbelt\ use) = \sum_{s=1}^{S} P(seat - belt\ use|C_s) \times P(C_s) \tag{4}$$

where $P(seatbelt\ use)$ is the unconditional probability of seatbelt use, $P(seatbelt\ use|C_s)$ is the conditional probability of seatbelt use in class C_s (same as Eq. 2), and $P(C_s)$ is the probability of class C_s. The overall log-likelihood function can be determined by the product of Eq. (4) over the entire observations. The Maximum Likelihood Estimation (MLE) approach is employed for the estimation of the latent class logit model.

2.3. Measures of goodness-of-fit

In the above formulation of the latent class model, the classes are assumed to be latent across observations, and thus the number of latent classes is not known a priori. Therefore, the model is empirically tested with a different number of classes (S), and the preferred number of classes is selected based on the model with the superior statistical fit. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are employed to compare the statistical fit of the model candidates (Washington et al., 2020):

$$AIC = -2LL + 2P \tag{5}$$

$$BIC = -2LL + PLog_{(N)} \tag{6}$$

where LL is the log-likelihood of the estimated model at convergence, P is the number of estimated parameters, and N is the number of observations or sample size. The model with lower AIC and BIC is regarded as a superior model in terms of statistical fit.

2.4. Marginal effects

To assess the effects of explanatory variables on the probability of seatbelt use, marginal effects are calculated as the amount of change in the probability of seatbelt use as a result of one unit change in a continuous explanatory variable (or a change from “0” to “1 “in a dummy variable) while holding all other explanatory variables at their means (Washington et al., 2020):

$$ME(x_i) = d\left(\frac{e^{\beta X_i}}{1 + e^{\beta X_i}}\right) \Big/ dx_i \tag{7}$$

Where $ME(x_i)$ is the marginal effect of variable x_i. For latent class models, marginal effects are the summation of marginal effects for each class weighted by their estimated latent class probabilities (Hensher et al., 2005).

3. Results and discussion

Several variants of the latent class logit model (Eq. 4) were first estimated against empirical data in this study using different numbers of classes, and their statistical fit was then compared to select the superior model. Explanatory variables were inserted into the models using stepwise variable selection criterion. In all models, explanatory variables were tested for multicollinearity by computing the Pearson correlation coefficients, and the variables with unacceptably high (>0.7) correlation coefficients were excluded from the models. The statistical fits of these models are presented in Table 3.

The results indicate that the logit model with three classes has lower AIC and BIC values (120252.2 and 120809.3, respectively) compared to the other variants of this model. This finding implies that the sample data used in this study consists of three latent classes to which vehicle occupants may belong with certain probabilities. As a result, the latent class logit model with three classes is now selected as the superior model for making inferences about the effects of the explanatory variables on seatbelt use. Table 4 presents the results of this model for the sample data in this study.

3.1. Class membership model component

According to Table 4, vehicle occupants may belong to the first, the second, and the third latent class with 10.2 %, 23.8 %, and 66.0 % probability, respectively. The average predicted seatbelt use rates in these classes are 81.1 %, 89.9 %, and 94.6 %, respectively. These predicted rates and their corresponding Gaussian densities are also shown in Fig. 1 and imply that the third class is associated with more conservative (conformist) and compliant drivers, whereas the first and the second classes are associated with more eccentric and less compliant drivers. Interestingly, the determinants of class probabilities paint a similar picture of these classes. Gender, average income per capita, and vehicle body type are the statistically significant factors determining the probabilities of the three latent classes. The parameters of these factors in Table 3 show that drivers in larger vehicles are more likely to fall within the first and the second classes (classes with lower seat belt use) in comparison with the third class whereas drivers coming from locations with higher income per capita are less likely to fall within these two classes. These findings are consistent with the previous studies in the literature in that larger vehicles and lower-income are all associated with lower seatbelt use (Houston and Richardson, 2005; Wells et al., 2002; Boyle et al., 2003; Glassbrenner and Ye, 2007, Gkritza and Mannering, 2008; CTR, 2018). The positive parameter of gender in the probability of the second class indicates that male drivers are more likely to fall in the second class (i.e., lower seatbelt use group) in comparison with the third class. This finding is also intuitive considering that male drivers have been reported to have lower seatbelt use (Houston and Richardson, 2005; Wells et al., 2002; Boyle et al., 2003; Gkritza and Mannering, 2008; Glassbrenner et al., 2004; Hezaveh and Cherry, 2019a,b; Hezaveh et al., 2019b).

3.2. Seatbelt use model component

The results of the seatbelt use component of the latent class model

Table 3
Model selection based on statistical measures of fit.

	Binary logit model	Latent class binary logit model with 2 classes	Latent class binary logit model with 4 classes	Latent class binary logit model with 3 classes
Likelihood at convergence	−60561.5	−60274.2	−60103.9	−60076.1
Number of parameters	15	32	64	53
Number of observations	242,468	242,468	242,468	242,468
AIC	121153.0	120412.5	120335.9	120252.2
BIC	121309.0	120945.2	121001.3	120809.3

Table 4
Results of the latent class binary logit model (with three classes) of drivers' seat belt use in single-occupant vehicles.

Variable	Latent class 1		Latent class 2		Latent class 3	
	Estimate	*t*-Statistic	Estimate	*t*-Statistic	Estimate	*t*-Statistic
Class membership model component						
Constant	−1.293	−5.70	−0.877	−5.92	[a]	[a]
Gender: male	–	–	0.615	6.60	[a]	[a]
Income per capita	−2.528	−3.30	−2.068	−8.66	[a]	[a]
Vehicle body type: large passenger vehicles	0.275	2.38	0.216	3.90	[a]	[a]
Class probabilities (population share)	0.102		0.238		0.660	
Average seatbelt use within each class	0.811		0.899		0.946	
Seatbelt use model component						
Constant	3.685	3.380	0.426	1.480	2.569	7.020
Age: less than 25 years old	–	–	–	–	−0.172	−2.680
Age: between 25 and 40 years old	–	–	–	–	−0.177	−2.720
Alcohol consumption	−6.330	−4.660	−5.113	−6.950	2.154	2.120
Distraction	−3.418	−6.680	0.541	2.590	–	–
Drug consumption	–	–	−3.342	−4.060	−1.545	−3.720
Gender: female	–	–	−0.749	−4.710	0.866	8.250
Total population density	–	–	−0.101	−2.470	0.243	3.750
Percentage of white race	−4.200	−3.870	3.072	10.820	−0.790	−2.040
Average travel time	9.446	4.930	–	–	–	–
Bachelor's education percentage	–	–	2.409	6.450	–	–
Income per capita	–	–	−5.985	−7.980	4.467	5.360
Day time	–	–	0.401	5.230	0.548	9.260
Weather – rainy	1.329	3.560	0.189	2.670	–	–

-: not statistically significant.
[a] Used as a reference in the class membership model component.

clearly show varied effects of explanatory variables (i.e., different sets of statistically significant variables and/or different parameter estimates for the same statistically significant variables) across the three classes of observations, indicating that the latent class logit model has been able to address unobserved heterogeneity in data.

Within the first class, the parameters of alcohol consumption and distraction are negative, indicating that alcohol consumption and distraction are associated with a decreased likelihood of seatbelt use. This finding is intuitive and consistent with the previous findings in the literature (Wilson, 1990). From the census tract information, the average percentage of the white race has a negative parameter, indicating that drivers coming from locations with a higher percentage of the white race are more likely to use the seatbelt. Briggs et al. (2006) also reported racial disparities in seatbelt use in the United States. The rainy weather condition has a positive parameter indicating that rainy weather is associated with a higher probability of seatbelt use. This finding is intuitive and implies that drivers are perhaps more cautious when driving in rainy weather conditions.

Within the second class, the parameters of alcohol consumption and drug consumption are negative, indicating that alcohol and drug consumption are associated with a decreased likelihood of seatbelt use. However, the distraction has a positive parameter in this class, indicating that distracted drivers are more likely to wear a seatbelt. This finding may be indicative of distracted drivers' self-awareness or perhaps overconfidence in their behaviour. In addition, female drivers and drivers coming from locations with higher income per capita are associated with a lower likelihood of seatbelt use in this class, which is also consistent with our earlier findings in that the second class is more represented by male drivers and drivers coming from locations with lower income per capita.

Interestingly, the average percentage of Bachelor's as the highest education level and the average percentage of the white race have positive parameters indicating that drivers who are coming from locations with higher education levels and a higher percentage of the white race are associated with an increased likelihood of seatbelt use. On the contrary, locations with a higher population density are associated with a decreased likelihood of seatbelt use, which is intuitive considering that these locations could consist of a higher proportion of eccentric individuals—which is the characteristics of this class. Finally, driving during the day and in rainy weather conditions have an increasing effect on the likelihood of seatbelt use.

Within the third class, drug consumptions is associated with a decreased likelihood of seatbelt use, similar to the second class. However, alcohol consumption has a positive parameter in this class, indicating that drivers with alcohol in their blood are more likely to wear a seatbelt. Considering that this class is associated with more conformist drivers, this finding may reflect the endogeneity between drivers' seatbelt use and their alcohol consumption in this class in that those drivers who are more conformist in wearing a seatbelt are more likely to consume alcohol because they are more confident and self-aware of their driving behaviour. The parameters of the two categories of age (less than 25 years old and between 25 and 40 years old) are negative, indicating that drivers younger than 40 years old are less likely to use a seatbelt. This finding is consistent with the previous findings in the literature

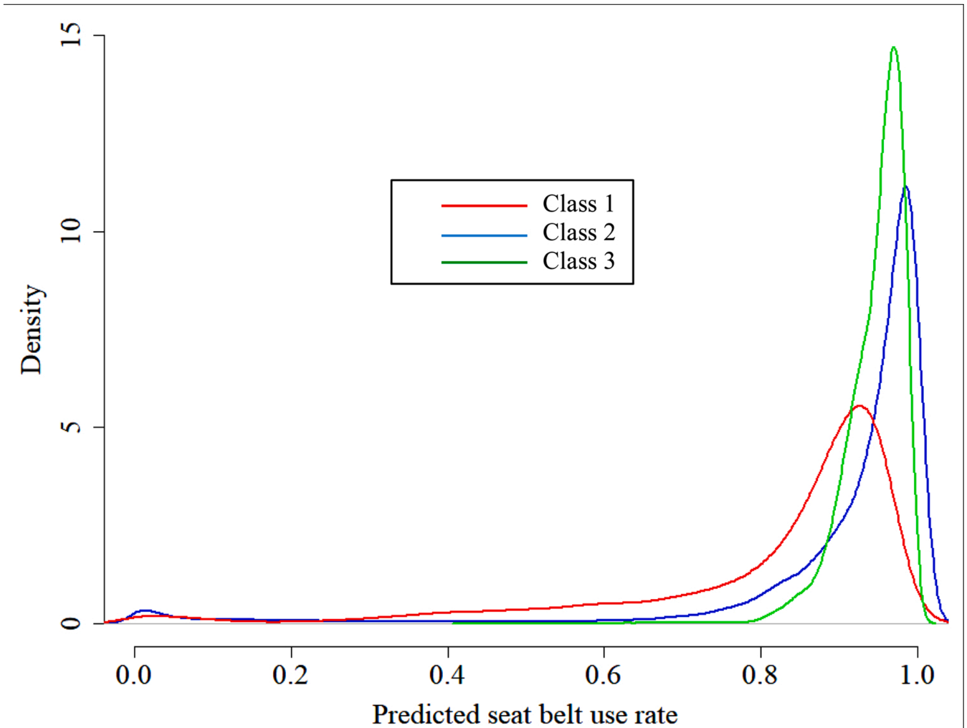

Fig. 1. Predicted density of drivers' seat belt use across latent classes in single-occupant vehicles.

(Glassbrenner et al., 2004; Calisir and Lehto, 2002). Moreover, female drivers and those drivers coming from locations with higher total population density and higher income per capita are more likely to wear a seatbelt in this class. Finally, driving in rainy weather conditions is associated with an increased likelihood of seatbelt use.

The parameters of the latent class logit model provided a comprehensive picture of the varied effects of external factors on seatbelt use within different possible populations of drivers. However, in order to obtain a more tangible understanding of the effects of these external factors on seatbelt use, their marginal effects were calculated for the overall logit model (Table 5).

According to the marginal effects, drivers with alcohol and drug in their blood have a 0.019 higher probability and 0.078 lower probability of wearing a seatbelt, respectively. In addition, male drivers have a 0.018 lower probability of wearing a seatbelt in comparison with female drivers. On the contrary, the probability of wearing a seatbelt during the day is 0.022 higher than during the night. In terms of census tract information, the marginal effects indicate that every additional minute in the average travel time is associated with 0.042 higher probability of wearing a seatbelt. Moreover, the probability of seatbelt use is 0.032 higher if the percentage of higher (Bachelor's) education level increases by 1 unit. Finally, every additional dollar of income per capita is associated with 0.077 higher probability of wearing a seatbelt. These findings show the great importance of policy investment in local neighbourhoods of Tennessee in order to increase seatbelt use rate, which in turn could lead to improving road safety.

Table 5
Marginal effects of variables in the latent class logit model of drivers' seat belt use in single-occupant vehicles.

Variable	Marginal effect
Age: less than 25 years old	−0.005
Age: between 25 and 40 years old	−0.007
Alcohol consumption	**−0.019**[a]
Distraction	**0.005**
Drug consumption	−0.078
Gender: male	−0.018
Total population density	0.005
Percentage of the white race	**−0.010**
Average travel time	0.042
Percentage of Bachelor's degree as the highest education	0.032
Income per capita	0.077
Day time	0.022
Weather – rainy	0.008

[a] Bold numbers indicate that the marginal effects are not statistically significant.

4. Conclusions

Seatbelt use enforcement is considered to be an effective strategy worldwide to reduce road injuries and fatalities. Despite the known benefits of seatbelt use, not wearing seatbelt still contributes to a substantial proportion of road fatalities in Tennessee, United States. As such, this study aimed to investigate the determinants of seatbelt use in single-occupant vehicles in this state.

Our findings indicate that the seatbelt use behaviour of drivers in single-occupant vehicles is rather heterogeneous in the state of Tennessee. While these drivers have a high seatbelt use rate in general, they are likely to belong to two general categories, conformist and eccentric, with higher seatbelt use rate in the former category and lower seatbelt use rate in the latter category. Male drivers, drivers of larger vehicles, and those drivers coming from lower-income neighbourhoods are more likely to belong to the eccentric category, whereas female drivers, drivers of passenger cars, and drivers coming from higher-income neighbourhoods are more likely to belong to the conformist category. This finding is important as the effects of a few external factors such as alcohol consumption on seatbelt use are different across these two categories. This finding may help policymakers better understand that their decisions may not have the same effects on seatbelt use for all drivers.

The results of the overall analysis indicate that male drivers, younger drivers, and drivers who have consumed drugs are less likely to wear a seatbelt, whereas drivers who come from areas with higher population density, higher education levels, and higher income per capita are more likely to wear a seatbelt. These findings may help policymakers implement decisions (for example, investing in increasing education levels and economic growth) at an aggregate level to increase seatbelt use rate and improve safety.

This study is not without limitations. We only focused on the seatbelt use of drivers in single-vehicle occupants in Tennessee. Investigating the seatbelt use of drivers and passengers in multi-occupant vehicles is a worthwhile research direction, particularly because of the interactions between seatbelt use of drivers and occupants in those vehicles. In addition, we did not consider the effects of unobserved heterogeneity within different classes of observations. Such heterogeneity may be significant across vehicle occupants especially because we allocated the same census tract information to all individuals coming from one census tract. However, this assumption may not be totally accurate. Employing appropriate methodological approaches such as random parameters models can capture the differences between individuals living in the same census tract.

Author statement

A.P. Afghari, A.M. Hezaveh, M.M. Haque: study conception and design; **A.M. Hezaveh, C. Cherry:** data collection; **A.P. Afghari, A.M. Hezaveh, M.M. Haque:** analysis and interpretation of results; **A.P. Afghari:** draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Competing Interest

The authors report no declarations of interest.

References

Abay, K.A., Paleti, R., Bhat, C.R., 2013. The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity. Transp. Res. Part B Methodol. 50, 74–89.

Abolhassani, L., Afghari, A.P., Borzadaran, H.M., 2019. Public preferences towards bicycle sharing system in developing countries: the case of Mashhad. Iran. Sustainable cities soc. 44, 763–773.

Afghari, A.P., Haque, M.M., Washington, S., Smyth, T., 2016. Bayesian latent class safety performance function for identifying motor vehicle crash black spots. Transp. Res. Rec. 2601 (1), 90–98.

Afghari, A.P., Washington, S., Haque, M.M., Li, Z., 2018. A comprehensive joint econometric model of motor vehicle crashes arising from multiple sources of risk. Anal. Methods Accid. Res. 18, 1–14.

Ali, M., Haidar, N., Ali, M.M., Maryam, A., 2011. Determinants of seat belt use among drivers in Sabzevar, Iran: a comparison of theory of planned behavior and health belief model. Traffic Inj. Prev. 12 (1), 104–109.

Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41 (1), 153–159.

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

Behnood, A., Mannering, F., 2017b. The effect of passengers on driver-injury severities in single-vehicle crashes: a random parameters heterogeneity-in-means approach. Anal. Methods Accid. Res. 14, 41–53.

Blincoe, L., Miller, T.R., Zaloshnja, E., Lawrence, B.A., 2015. The Economic and Societal Impact of Motor Vehicle Crashes, 2010 Revised. National Highway Traffic Safety Administration, Washington, DC. Report No. DOT HS 812 013.

Boyle, J.M., Vanderwolf, P., Schulman, R., 2003. 2003 motor Vehicle Occupant Safety Survey, 2. National Highway Traffic Safety Administration, United States. Safety belt report. Report No. DOT-HS-809-789.

Briggs, N.C., Schlundt, D.G., Levine, R.S., Goldzweig, I.A., Stinson Jr., N., Warren, R.C., 2006. Seat belt law enforcement and racial disparities in seat belt use. Am. J. Prev. Med. 31 (2), 135–141.

Buddhavarapu, P., Scott, J.G., Prozzi, J.A., 2016. Modeling unobserved heterogeneity using finite mixture random parameters for spatially correlated discrete count data. Transp. Res. Part B Methodol. 91, 492–510.

Calisir, F., Lehto, M.R., 2002. Young drivers' decision making and safety belt use. Accid. Anal. Prev. 34 (6), 793–805.

Chaudhary, N.K., Preusser, D.F., 2006. Connecticut nighttime safety belt use. J. Safety Res. 37 (4), 353–358.

Chaudhary, N.K., Alonge, M., Preusser, D.F., 2005. Evaluation of the Reading, PA nighttime safety belt enforcement campaign: september 2004. J. Safety Res. 36 (4), 321–326.

Cherry, C.R., Hezaveh, A.M., Noltenius, M., Khattak, A., Merlin, L., Dumbaugh, E., Sandt, L., 2018. Completing the Picture of Traffic Injuries: Understanding Data Needs and Opportunities For Road Safety. University of Tennessee, Knoxville.

CTR, 2018. 2017 Survey Of Safety Belt Usage In Tennessee Final Report. Center for Transportation Research, University of Tennessee, Knowxville, United States.

Dee, T.S., 1998. Reconsidering the effects of seat belt laws and their enforcement status. Accid. Anal. Prev. 30 (1), 1–10.

Eby, D.W., Molnar, L.J., Olk, M.L., 2000. Trends in driver and front-right passenger safety belt use in Michigan: 1984–1998. Accid. Anal. Prev. 32 (6), 837–843.

Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Accid. Anal. Prev. 39 (5), 1037–1049.

Foster, S.A., 1999. The geography of behaviour: an evolutionary perspective. Trends Ecol. Evol. 14 (5), 190–195.

Fountas, G., Anastasopoulos, P.C., Abdel-Aty, M., 2018. Analysis of accident injury-severities using a correlated random parameters ordered probit approach with time variant covariates. Anal. Methods Accid. Res. 18, 57–68.

Gkritza, K., Mannering, F.L., 2008. Mixed logit analysis of safety-belt use in single-and multi-occupant vehicles. Accid. Anal. Prev. 40 (2), 443–451.

Glassbrenner, D., Carra, J.S., Nichols, J., 2004. Recent estimates of safety belt use. J. Safety Res. 35 (2), 237–244.

Glassbrenner, D., & Ye, T. J. (2007). Seat Belt Use in 2007–overall Results (pp. 1-4). Washington, DC: National Highway Traffic Safety Administration.

Greene, W.H., Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transp. Res. Part B Methodol. 37 (8), 681–698.

Hamed, M.M., Al-Eideh, B.M., 2020. An exploratory analysis of traffic accidents and vehicle ownership decisions using a random parameters logit model with heterogeneity in means. Anal. Methods Accid. Res. 25, 100116.

Hensher, D.A., Rose, J.M., Greene, W.H., 2005. Applied Choice Analysis: a Primer. Cambridge university press.

Heydari, S., Fu, L., Thakali, L., Joseph, L., 2018. Benchmarking regions using a heteroskedastic grouped random parameters model with heterogeneity in mean and variance: applications to grade crossing safety analysis. Anal. Methods Accid. Res. 19, 33–48.

Hezaveh, A.M., Cherry, C.R., 2019a. Neighborhood-level factors affecting seat belt use. Accid. Anal. Prev. 122, 153–161.

Hezaveh, A.M., Cherry, C.R., 2019b. Neighborhood-level factors affecting seat belt use. Accid. Anal. Prev. 122, 153–161.

Hezaveh, A.M., Zavareh, M.F., Cherry, C.R., Nordfjærn, T., 2017. Errors and violations in relation to bicyclists' crash risks: development of the Bicycle Rider Behavior Questionnaire (BRBQ). J. Transp. Health 8, 289–298.

Hezaveh, A.M., Nordfjærn, T., Mamdoohi, A.R., Şimşekoğlu, Ö., 2018. Predictors of self-reported crashes among iranian drivers: exploratory analysis of an extended driver behavior questionnaire. PROMET-Traffic Transportation 30 (1), 35–43.

Hezaveh, A.M., Arvin, R., Cherry, C.R., 2019a. A geographically weighted regression to estimate the comprehensive cost of traffic crashes at a zonal level. Accid. Anal. Prev. 131, 15–24.

Hezaveh, A.M., Nordfjærn, T., Everett, J., Cherry, C.R., 2019b. The correlation between education, engineering, enforcement, and self-reported seat belt use in Tennessee: incorporating heterogeneity and time of day effects. Transp. Res. Part F Traffic Psychol. Behav. 66, 379–392.

Hodson-Walker, N., 1970. The value of safety belts: a review. Can. Med. Assoc. J. 102 (4), 391.

Houston, D.J., Richardson, L.E., 2005. Getting Americans to buckle up: the efficacy of state seat belt laws. Accid. Anal. Prev. 37 (6), 1114–1120.

Huang, H., Chang, F., Zhou, H., Lee, J., 2019. Modeling unobserved heterogeneity for zonal crash frequencies: a Bayesian multivariate random-parameters model with mixture components for spatially correlated data. Anal. Methods Accid. Res. 24, 100105.

Kadilar, G.O., 2016. Effect of driver, roadway, collision, and vehicle characteristics on crash severity: a conditional logistic regression approach. Int. J. Inj. Contr. Saf. Promot. 23 (2), 135–144.

Kamruzzaman, M., Hine, J., 2013. Self-proxy agreement and weekly school travel behaviour in a sectarian divided society. J. Transp. Geogr. 29, 74–85.

Knapper, C., Cropley, A.J., Moore, R., 1976. Attitudinal factors in the non-use of seat belts. Accid. Anal. Prev. 8 (4), 241–246.

Lund, A.K., 1986. Voluntary seat belt use among US drivers: geographic, socioeconomic and demographic variation. Accid. Anal. Prev. 18 (1), 43–50.

Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.

Martinussen, L.M., Hakamies-Blomqvist, L., Møller, M., Özkan, T., Lajunen, T., 2013. Age, gender, mileage and the DBQ: the validity of the Driver Behavior Questionnaire in different driver groups. Accid. Anal. Prev. 52, 228–236.

McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, D., McFadden (Eds.), A Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge, MA.

Merlin, L.A., Cherry, C.R., Mohamadi-Hezaveh, A., Dumbaugh, E., 2020. Residential accessibility's relationships with crash rates per capita. J. Transp. Land Use 13 (1), 113–128.

Morgan, M.A., 2015. Analyzing seat belt usage from accident data: an evaluation of click it or Ticket enforcement campaigns in Ohio. Int. J. Police Sci. Manag. 17 (1), 32–39.

NHTSA, 2009. Nigttime Enforcement of Seat Belt Laws: An Evaluation of Three Community Programs. National Highway Traffic Safety Administration. Report No. DOT HS, 811, 189.

NHTSA, 2017. Seat Belt Use in 2017—Use Rates in the States and Territories. In Analysis, N.S.N.C.F.S.A. ed..

Nichols, J.L., Tison, J., Solomon, M.G., Ledingham, K.A., Preusser, D.F., Preusser Research Group, 2009. Evaluation of a Rural Demonstration Program to Increase Seat Belt Use in the Great Lakes Region. National Highway Traffic Safety

Administration. Office of Behavioral Safety Research, United States (No. DOT HS 811 084).
Nordfjærn, T., Hezaveh, A.M., Mamdoohi, A.R., 2015. An analysis of reported driver behaviour in samples of domestic and expatriate Iranians. J. Risk Res. 18 (5), 566–580.
Office Of Energy Efficiency & Renewable Energy, 2018. Average Vehicle Occupancy Remains Unchanged From 2009 to 2017.
Okamura, K., Fujita, G., Kihira, M., Kosuge, R., Mitsui, T., 2012. Predicting motivational determinants of seatbelt non-use in the front seat: a field study. Transp. Res. Part F Traffic Psychol. Behav. 15 (5), 502–513.
Park, B.J., Lord, D., 2009. Application of finite mixture models for vehicle crash data analysis. Accid. Anal. Prev. 41 (4), 683–691.
Pickrell, T., Ye, T., 2009. Traffic Safety Facts (Research Note): Seat Belt Use in 2008–Demographic Results. National Highway Traffic Safety Administration/ Department of Transportation, Washington DC.
Reagan, I.J., McClafferty, J.A., Berlin, S.P., Hankey, J.M., 2013. Using naturalistic driving data to identify variables associated with infrequent, occasional, and consistent seat belt use. Accid. Anal. Prev. 50, 600–607.
Reinfurt, D.W., 2004. Documenting the sustainability of a mature click it or Ticket program: the North Carolina experience. J. Safety Res. 35 (2), 181–188.
Rentfrow, P.J., 2010. Statewide differences in personality: toward a psychological geography of the United States. Am. Psychol. 65 (6), 548.
Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., Milton, J., 2017. Occupant injury severities in hybrid-vehicle involved crashes: a random parameters approach with heterogeneity in means and variances. Anal. Methods Accid. Res. 15, 41–55.
Shakya, I., Shults, R.A., Stevens, M.R., Beck, L.F., Sleet, D.A., 2020. State-level seat belt use in the United States, 2011–2016: comparison of self-reported with observed use and use by fatally injured occupants. J. Safety Res. 73, 103–109.
Shaon, M.R.R., Qin, X., Afghari, A.P., Washington, S., Haque, M.M., 2019. Incorporating behavioral variables into crash count prediction by severity: a multivariate multiple risk source approach. Accid. Anal. Prev. 129, 277–288.
Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., Haque, M., 2020. Is an informed driver a better decision maker? A grouped random parameters with heterogeneity-in-means approach to investigate the impact of the connected environment on driving behaviour in safety-critical situations. Anal. Methods Accid. Res., 100127
Solomon, M.G., Compton, R.P., Preusser, D.F., 2004. Taking the click it or ticket model nationwide. J. Safety Res. 35 (2), 197–201.
Solomon, M.G., Chaudhary, N.K., Preusser, D.F., Preusser Research Group, 2007. Daytime and Nighttime Seat Belt Use at Selected Sites in New Mexico. Report No. DOT-HS-810-705. National Highway Traffic Safety Administration, United States.
Thomas, F.D., Blomberg, R.D., Peck, R.C., Cosgrove, L.A., Salzberg, P.M., 2008. Evaluation of a high visibility enforcement project focused on passenger vehicles interacting with commercial vehicles. J. Safety Res. 39 (5), 459–468.
Thomas, A.M., Cook, L.J., Olson, L.M., 2011. Evaluation of the click it or ticket intervention in Utah. Accid. Anal. Prev. 43 (1), 272–275.
Thomas, F.D., Blomberg, R.D., Masten, S.V., Peck, R.C., Van Dyk, J., Cosgrove, L.A., 2017. Evaluation of the Washington Nighttime Seat Belt Enforcement Program. Department of Transportation. National Highway Traffic Safety Administration, United States. Report No. DOT HS 812 395.
Tison, J., Williams, A.F., 2010. Analyzing the First Years of the Click It or Ticket Mobilizations. Report No. DOT HS 811 232. National Highway Traffic Safety Administration, United States.
Tison, J., Williams, A.F., Chaudhary, N.K., 2010. Daytime and Nighttime Seat Belt Use by Fatally Injured Passenger Vehicle Occupants. National Highway Traffic Safety Administration, United States. Report No. DOT HS 811 281.
Van Acker, V., Van Wee, B., Witlox, F., 2010. When transport geography meets social psychology: toward a conceptual model of travel behaviour. Transp. Rev. 30 (2), 219–240.
Vivoda, J.M., Eby, D.W., Kostyniuk, L.P., 2004. Differences in safety belt use by race. Accid. Anal. Prev. 36 (6), 1105–1109.
Vivoda, J.M., Eby, D.W., St. Louis, R.M., Kostyniuk, L.P., 2007. A direct observation study of nighttime safety belt use in Indiana. J. Safety Res. 38 (4), 423–429.
Waseem, M., Ahmed, A., Saeed, T.U., 2019. Factors affecting motorcyclists' injury severities: an empirical assessment using random parameters logit model with heterogeneity in means and variances. Accid. Anal. Prev. 123, 12–19.
Washington, S., Karlaftis, M., Mannering, F., Anastasopoulos, P., 2020. Statistical and Econometric Methods for Transportation Data Analysis, third edition. CRC Press, Taylor and Francis Group, New York, NY.
Wells, J.K., Williams, A.F., Farmer, C.M., 2002. Seat belt use among African Americans, Hispanics, and whites. Accid. Anal. Prev. 34 (4), 523–529.
Wilson, R., 1990. The relationship of seat belt non-use to personality, lifestyle and driving record. Health Educ. Res. 5 (2), 175–185.
Xiong, Y., Mannering, F.L., 2013. The heterogeneous effects of guardian supervision on adolescent driver-injury severities: a finite-mixture random-parameters approach. Transp. Res. Part B Methodol. 49, 39–54.