Article

The Latent Class Multinomial Logit Model for Modeling Front-Seat Passenger Seatbelt Choice, Considering Seatbelt Status of Driver

Mahdi Rezapour * and Khaled Ksaibati

Wyoming Technology Transfer Center, University of Wyoming, 1000 E. University Avenue, Department 3295, Laramie, WY 82071, USA; khaled@uwyo.edu

*   Correspondence: mrezapou@uwyo.edu; Tel.: +1-307-766-6230; Fax: +1-307-766-6784

Abstract: The literature review highlighted the impacts of drivers' behavior on passengers' attitudes in the choice of seatbelt usage. However, limited studies have been done to determine those impacts. Studying the passengers' seatbelt use is especially needed to find out why passengers choose not to buckle up, and consequently it helps decision makers to target appropriate groups. So, this study was conducted to find drivers' characteristics that might impact the passenger's seatbelt use, in addition to other passengers' characteristics themselves. While performing any analysis, it is important to use a right statistical model to achieve a less biased point estimate of the model parameters. The latent class multinomial logit model (LC-MNL) can be seen as an alternative to the mixed logit model, replacing the continuous with a discrete distribution, by capturing possible heterogeneity through membership in various clusters. In this study, instead of a response to the survey or crash observations, we employed a real-life observational data for the analysis. Results derived from the analysis reveal a clear indication of heterogeneity across individuals for almost all parameters. Various socio-demographic variables for class allocation and models with different latent numbers were considered and checked in terms of goodness of fit. The results indicated that a class membership with three factors based on vehicle type would result in a best fit. The results also highlighted the significant impacts of driver seatbelt status, time of a day, distance of traveling, vehicle type, and driver gender, instead of passenger gender, as some of the factors impacting the passengers' choice of seatbelt usage. In addition, it was found that the belting status of passengers is positively associated with the belting condition of drivers, highlighting the psychological behavioral impact of drivers on passengers. Extensive discussion has been made regarding the implications of the findings.

Keywords: preference heterogeneity; passenger seatbelt; passenger behavior; traffic safety; choice of wearing seatbelt

1. Introduction

More than 37,000 people died in highway crashes in the U.S. in 2017 alone, out of which 47% vehicle occupants were not wearing seatbelt [1]. It has been noted that seatbelt usage could reduce both fatal and non-fatal injuries by 60% among front seat occupants, and 44% for rear seat occupants [2]. Despite efforts in regard to the importance of seatbelt usage, a significant portion of vehicles users still do not buckle up. The number is much higher for passengers of vehicles compared to the drivers. For instance, based on our data while in Wyoming more than 80% of the drivers buckle up, less than 50% of the passenger's used seatbelt. Identifying factors impacting the choice of wearing seatbelt is a first step in efforts to increase the rate of buckled occupants.

While performing statistical analysis for passengers' seatbelt use, a few research questions are worthy of investigation. How does the passenger view the safety of driver, and the vehicle they are traveling in? Do they care about riding with a safe driver?

This study is set up to answer those specific questions, in addition to other unanswered questions.

The next few paragraphs highlight several studies conducted on the choice of wearing a seatbelt. In a recent study, self-reported seatbelt usage for adults in front and rear passenger seats was evaluated [3]. The results showed that seatbelt usage for the front passenger is higher (86%) than the rear passengers (62%). In another study, a telephone survey was conducted in 2016 targeting passengers who had recently ridden in the rear seat and did not always use a seatbelt [4]. The results indicated that seatbelt use is lower among passengers who travel in the rear seat, and the reasons for not buckling up included forgetfulness, inconvenience, and discomfort.

Driver and passenger seatbelt use among US high school students were evaluated [5]. Based on the descriptive statistics it was found that 59% of students use the seatbelt while driving, only 42% of passengers are always buckled up. In a similar study, seatbelt use across high school students in the US by means of summary statistics was evaluated [6]. The results highlighted that the younger drivers have lower seatbelt use, and passengers have lower seatbelt use while riding with teenage drivers compared with adults.

The relationship between driver and passenger seatbelt use was investigated by bivariate Probit model. The data were collected by means of survey. A positive correlation was found between the rate of drivers and passengers seatbelt use [7]. In another study, the determinant of seatbelt use by means of FARS data was investigated [8]. The results highlighted that the passengers are less likely to be buckled compared with drivers.

Associated factors among passengers were investigated in another study [9]. The association was investigated by means of a motor vehicle crash (MVC) dataset. The results of descriptive analysis highlighted that there is a positive correlation between the seatbelt use of drivers and passengers.

The seat belt use in Wisconsin was used to investigate the relationship between seatbelt use of drivers and passengers [10]. The descriptive analysis of the data highlighted that there is a significant difference between seatbelt use of drivers and passengers. Moreover, it was highlighted that there is a positive correlation between seatbelt use of drivers and passengers.

Despite efforts, not many studies conducted using non-survey data for investigating the factors to the passenger's seatbelt use. Additionally, the literature review lacks a reliable statistical method for finding unbiased factors concerning the choice of passengers to use a seatbelt. Using appropriate statistical methods could help to explain the underlying causes of the choice being made by drivers and passengers not to buckle up.

In statistical analyses, while some researchers assume homogenous preference across individuals, others account for heterogeneity by using a continuous distribution for random parameters, allowing the models' parameters to vary randomly across the individuals [11], or using a discrete distribution by stratifying the individuals into predefined numbers of classes. Accounting for possible heterogeneity would expand the understanding regarding the real impacts of various factors on the response. To account for heterogeneity across observations in the dataset, the latent class multinomial logit model (LC-MNL) was used.

Study Contributions

The contributions of this study are summarized as follows:

1. The majority of past studies used a response to the survey or crash observations as their datasets. However, it is important to gain a better understanding by using a real-life information. So, in this study, we employed a dataset derived from real observations.
2. The majority of past studies failed to account for the heterogeneity in the dataset and the analysis was undertaken by means of solely traditional descriptive or statistical analyses. However, it is important to account for the observations' heterogeneity due to variation across drivers and passengers. Failure to account for the heterogeneity is likely to result in biased or misguided results. So, this study was conducted to

provide a more reliable understanding regarding factors affecting the seatbelt use of passengers while considering the belting status of drivers.

3. While investigating factors affecting the choice of passenger seatbelt use, it is important to take into consideration all drivers and passengers characteristics, so here we took into consideration all those factors in the analysis. That is especially important as it is expected that behaviors of drivers to impact the passengers and vice versa.

This is one of the earliest studies using real-life dataset, while accounting for the dataset heterogeneity. Moreover, one of the earliest studies used a real-life dataset to capture factors to the seatbelt use of passengers while considering the belting status of drivers. The rest of this paper is structured as follows. Section 2 discusses the methodological formulation of the study in detail, while Section 3 provides an overview of the variable characteristics used in this study and Sections 4 and 5 outline and discuss the results.

2. Method

This method section is presented in two subsections. First, a general description about the latent class LC-MNL would be given. This is followed by an explanation regarding the way that the class allocation would be made based on various socio-econometric characteristics.

2.1. Latent Class Model

The idea behind the latent class model is that the data could be divided into S homogenous distinct classes and each class has its own parameters. The latent class, compared with mixed model, is less restrictive as it identifies classes without distributional assumptions. The latent class seeks unobserved grouping in the data. The parameters for first class should be normalized to zero for allowing of the segment's identification. The multinomial logit model (MNL) is a general form of LC-MNL, and its probability could be written as [12]:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}} \tag{1}$$

where P_{ni} is the probability of a passenger n making choice of bucking up or not as i.

Equation (1) could be extended to the mixed or LC model. An alternative to the mixed model, with continuous distribution is to allow a finite mixture of values for elements in β, with its associated probability [13]. Allowing β with k different elements, the log-likelihood would be written as:

$$\mathcal{LL} = \sum_{N=1}^{N} ln \sum_{s=1}^{S} \pi_s \prod (P_{ni|s})^{y_{ni}} \tag{2}$$

where π_s is the probability of an individual belonging to a class s, where $\sum_{s=1}^{S} \pi_s = 1$ and $0 \leq \pi_s \leq 1$, and P_{ni} was defined earlier, and would be estimated based on Equation (1). N is number of observations, and π_s are the class probabilities, which could be written as [14]:

$$\pi_s = \frac{exp(h^{\intercal}\gamma_s)}{\sum_{s=1}^{S} h^{\intercal}\gamma_s} \tag{3}$$

where h are a set of various explanatory variables being used for class allocations and parameters for the first class, γ_1, would be set as zero. Heterogeneity in taste could be accommodated by assuming separate classes with different values for various coefficients. Thus, by considering S classes across the population, we would have S numbers of various coefficients $\beta_1, \ldots, \beta_S$, where the number of S classes would be specified based on a better fit. Unlike the mixed model where function is a parametric and continuous distribution like normal, (LC-MNL) assumes that the choice preferences are distributed based on some discrete distribution [15]. Thus, in LC-MNL, the unobserved heterogeneity is accompanied by a discrete number of S classes or segments of drivers with various preferences within each class.

Based on *S*, passengers of the same class, shares homogenous preference within a class. An optimal number of *S* in the latent class could be achieved by lowest Akaike information criterion (AIC), and Bayesian information criterion (BIC) value. The theory of the LC states that individual behaviors vary based on observable attributes and latent heterogeneity, which depends on factors that are unobserved by the analyst [14].

2.2. LC Segment Allocation

For latent class assignment, a constant value of various socio-econometric could be used as a means of class membership allocation. A constant assignment could be used instead of socio-econometric characteristics for class assignment [16], when the latent assignment is not dependent on any socio-demographic variable. However, the shortcoming here is that the individuals that are identical based on the included predictor would be considered as having identical sensitivities.

By including various socio-demographic for class allocation in a latent class model, a further insight could be obtained regarding the likely class allocation of a given individual on the sample distribution [13]. The whole range of available socio-demographic variables were considered in this study, and the models were compared in terms of goodness of fit.

After trying all possible options, the socio-demographic variable of vehicle type was found to improve the model performance compared with other values. After considering various components and heterogeneity for latent allocations, a best fit model was identified based on the Akaike information criterion (AIC), and Bayesian information criterion (BIC). These are valid fit measures as they penalize for the number of included predictors. This is fundamental as the number of parameters is an important issue for models with various numbers of latent.

3. Case Study and Data

The dataset collection was performed in 2019 across 17 counties, at 289 locations in Wyoming. The sample of counties and sites were developed based on the federal guideline being setup in 2012 by the WYDOT and served as the baseline for yearly survey. The data includes information for all vehicles occupants, including drivers and outboard passengers. Observers collected the data between June 3rd and June 9th of 2019 [17]. They collected about 25,000 observations of drivers and passengers in about 18,000 vehicles. The observers who collected the data were trained in a classroom before any data collection. This was done to conform to the criteria highlighted by the state observational seatbelt survey issued in 2011 by the National Highway Traffic Safety Administration. The survey followed the uniform criteria for the state observational survey of seatbelt use, 23 CFR [18].

The monitors were employed to even attend training sessions with observers, and to receive additional training separate from the observers. The quality control monitoring session consist of an extensive review of the directions applied to the monitors. Once the observers were in the field, to ensure the reliability of the observations for different observers, the quality-control monitors conducted random spot checks. The random site selections were determined during that session for reliability spot checks where monitoring would take place.

Following the classroom training, the observers took part in pilot test to assure the required skills and measure the accuracy of the observers. The selection of random site was determined to fulfill the reliability spot checks where monitoring occurs. The sites were adjusted to ensure the observations quality. The preselected sites were approved by the FHWA to comply with the uniform criteria for the state observational survey of seatbelt use.

As the objective of this study was to evaluate front-seat passenger seatbelt usage, only observations where the vehicle had a single passenger were considered. After removing a single driver with no passenger, and drivers with more than one passenger due to various passenger seatbelt status, the total number of dataset observations were reduced to 6533.

In the dataset, there was an indicator regarding the passengers' seatbelt, in addition to the drivers' seatbelt status.

Moreover, various driver and passenger characteristics were considered in this study. The consideration of both drivers and passengers' characteristics could help to unlock important information regarding the interactions of drivers and passengers regarding passengers' seatbelt use. A summary of the attributes, and levels of various contributory factors, along with related descriptive summaries of the important predictors are presented in Table 1. As can be seen from Table 1, most of the vehicle passengers were unbuckled while the reverse applies to the drivers' seatbelt status. Driver gender was included in Table 1 as it was found that this predictor is important in predicting the passengers' seatbelt status. The majority of vehicles that had a passenger on board were vehicles with non-Wyoming plates.

Table 1. Descriptive summary of important predictors.

Variables	Mean	Variance	Min	Max
Passenger seat, not belted versus belted *	0.68	0.218	0	1
Type of vehicle				
Pickup truck, pickup versus other *	0.33	0.221	0	1
SUV, SUV versus others *	0.07	0.068	0	1
Vehicle license, non-Wyoming residence versus others *	0.60	0.600	0	1
Environmental characteristics				
Sunny, sunny versus other *	0.68	0.217	0	1
Day of a week, weekdays versus weekend *	0.80	0.161	0	1
Driver characteristics				
Driver belt condition, belted versus not belted *	0.14	0.121	0	1
Driver gender, female versus male *	0.26	0.192	0	1
Environmental condition				
Time of data collection: 7:30–9:30 AM versus other *	0.16	0.133	0	1
Number of lanes, 2 lanes, a single lane *	0.41	0.243	0	1
Converse County versus others *	0.015	0.014	0	1

* Reference category as 0.

Dummy variables of various counties were considered and only Converse County was found to be important, being included in Table 1.

4. Results

The latent class models were estimated using two different specifications, namely by considering various numbers of latent components and socio-demographic covariates for class allocation. As discussed, based on various goodness of fits criteria, the best specification was three components with a vehicle type variable for class allocation. Table 2 summarizes the results of the finalist models, considering the three-component model with a covariate of vehicle type, as well as a constant for class allocation.

Table 2. The results of the latent class model buckled passenger as reference.

	Estimate	Std. Error	*p*-Value	Estimate	Std. Error	*p*-Value
	Model with 3 latent classes, Changing Class Allocation Based on Vehicle Type			Standard Model with 3 classes		
Class I						
intercept	16.76	140.516	0.91	18.58	311.690	0.95
pickup truck	−2.79	0.783	<0.005	−0.61	0.467	0.19
Day of a week	−1.71	0.436	<0.005	−1.456	0.365	<0.005
driver Belting status	−2.39	0.819	<0.005	−2.18	0.576	<0.005
driver Gender	−19.79	140.518	0.89	−18.934	311.693	0.95
vehicle license	5.05	1.517	<0.005	2.25	0.599	<0.005
vehicle type: SUV	−2.40	1.199	0.05	−1.29	0.588	0.03
Time: 7:30–9:30	1.12	0.572	0.05	1.99	0.589	<0.005
Number of lanes	−0.23	0.347	0.52	−0.389	0.262	0.14
weather condition: sunny	0.22	0.337	0.51	0.36	0.310	0.24
County, Converse	−5.23	2.142	0.01	0.39	0.898	0.66
Class II						
Intercept	44.11	95.318	0.64	30.989	36.548	0.40
pickup truck	−15.44	36.643	0.67	−21.54	25.729	0.40
Day of a week	−1.59	0.743	0.03	−0.30	0.531	0.57
driver Belting status	−2.31	3.851	0.55	6.49	10.383	0.53
driver Gender	−0.03	3.014	0.99	−0.90	0.722	0.21
vehicle license	−28.40	58.335	0.63	−15.54	18.228	0.39
vehicle type: SUV	−18.89	53.350	0.72	−0.87	0.964	0.37
Time: 7:30–9:30	−1.88	1.9510	0.34	−0.312	1.378	0.82
Number of lanes	0.04	0.699	0.96	0.64	0.463	0.17
Weather condition, sunny	13.56	26.128	0.60	1.01	0.495	0.04
County, Converse	11.19	24.200	0.64	10.978	15.378	0.48
Class III						
Intercept	5.14	2.0544	0.01	−8.63	90.604	0.92
pickup truck	−3.54	0.876	<0.005	−3.68	0.915	<0.005
Day of a week	−0.90	0.490	0.07	−0.68	0.560	0.22
driver Belting status	−0.36	0.487	0.46	−0.53	0.6464	0.41
driver Gender	1.31	0.756	0.08	13.156	90.589	0.88
vehicle license	−0.54	0.601	0.37	0.43	0.669	0.52
vehicle type: SUV	−1.05	1.344	0.44	−1.523	0.918	0.10
Time: 7:30–9:30	−1.79	0.470	<0.005	−3.40	0.878	<0.005
Number of lanes	0.44	0.313	0.16	0.98	0.450	0.05
Weather condition, sunny	−1.02	0.450	0.02	−0.90	0.65	0.17

Table 2. *Cont.*

	Estimate	Std. Error	*p*-Value	Estimate	Std. Error	*p*-Value
County, Converse	−0.92	0.667	0.17	−4.75	1.266	<0.005
(class)2, γ_2	−1.14	0.124	<0.005	−0.85	0.187	<0.005
(class)3, γ_3	0.16	0.1223	0.18	0.08	0.085	0.34
SUV.class2	0.47	0.263	0.07	–	–	–
SUV.class3	0.28	0.198	0.16	–	–	–
	Log Likelihood: −3547 AIC: 7167 BIC: 7417			Log Likelihood: −3549 AIC: 7203 BIC: 7359		

While AIC of three-segment model considering a socio-demographic variable of vehicle type, for class allocation, has a lower AIC, compared with a three-segment model with a constant value, the latter model has a lower BIC. However, as the class allocation of vehicle for class 2, considering class allocations with a covariate, is significant, and more significant variables could be observed for this model, this model is chosen over the other model. Additional discussions about the selection of the best fit model will be presented in the next subsection.

It should be noted that buckled passenger was considered as reference. From Table 2 only the second-class membership (γ_2) is dependent on a constant and vehicle type variable. As discussed earlier, the results of class 1 membership would not be presented in the Table 2 as the values for this class were set as zero. As can be seen from Table 2, the signs of variable are varied across various segments, suggesting that the effects are characterized by heterogeneity. For instance, while the coefficient of time of data-collection, 7:30–9:30 a.m., has a positive sign of 1.12 for the first component, this value has a negative coefficient for the third component, −1.79. It should also be noted that the impact is not important for the second component. The heterogeneities could also be observed across various segments in terms of significance. For example, the impact of driver seatbelt use is only significant for the first component.

Although day of week is an important predictor across the three segments, it was found that the impact of this variable on passenger seatbelt usage is more influential for the first segment compared with the second and the third ones. The result indicates that getting a ride as a passenger during weekdays decreases the likelihood of being unbuckled. The impact might be resulted from the fact that passengers are less strict while going out for pleasure compared with other times, or weekdays.

Driver seatbelt status was another important predictor that was found to impact the seatbelt status of passengers for only one class. Although heterogeneity exists for this predictor, it is worth discussing the impact. The results indicated that having an unbuckled driver motivates the passenger to be unbuckled. The results are in line with the results obtained by [19] showing that there is a positive correlation between driver status of seatbelt use and passenger. However, it should be noted that the data used in that study included statewide injury surveillance system involved in crashes. Also, only a Chi-square test was implemented in that study. However, the heterogeneity of this variable should be taken into consideration. The results also found that among all the counties in the state, County of Converse's respondents were found to have a higher likelihood of being buckled.

Although the impact of passenger gender on the seatbelt status of passenger itself was not important, the results indicated that driver gender has an important impact on the seatbelt status of the passenger. Again, due to heterogeneity across the observation, the impact is only significant for the third class. The results show that the probability of being unbuckled increases for the passenger in the third segment as the gender of the driver changes from male to female.

Here, we also incorporated various vehicle types and their license types as those factors are likely to provide unobserved information regarding various characteristics of drivers and possibly passengers. Due to associated heterogeneity across the observations, the impact of drivers' vehicles residency was found to be important only for the first component.

The results indicated that the odds of being unbuckled increases for passengers as the vehicle plate changes from Wyoming to non-Wyoming. This result might again be due the fact that while travelling for a longer distance, passengers might feel less obliged to be buckled. It should be noted that the impact has the highest magnitude across all the incorporated parameters.

Of all types of vehicles, the SUV was found to impact the likelihood of seatbelt usage for only the first segment. Moreover, that variable was found to impact the class allocation of the model, which resulted in a model fit enhancement. The last parameters are the time of a day, 7:30–9:30 a.m., which decreased the likelihood of being unbuckled, weather condition of sunny, which increase the likelihood of being unbuckled, and Converse County. Again, it should be noted that the inclusion of time of day or weather condition could provide information regarding unobserved characteristics of vehicle occupants that were not recorded at the time of data collection. For example, time of day 7:30–9:30 could provide information regarding carriers or the purpose of travel for vehicle occupants, e.g., being students.

4.1. Comparison across Various LCM

It is worth discussing how we came up with models with different number of latent classes and variables for class allocation. The socio-econometric covariate, along with a constant value were considered as determinants of the membership probability. AIC and BIC were used for finding the best fit model. Models with two and three latent classes were considered. Moreover, only covariates of driver seatbelt usage, day of a week, and vehicle type of SUV were considered as other covariates were found not to impact the class membership probability significantly.

Table 3 presents the summary of various goodness of fit measures being used for a comparison. While log likelihood does not penalize for the number of included predictors, AIC and BIC are reliable measures as they penalize for the number of parameters. In finding an optimal number of segments, both two and three were assessed. The results support the existence of heterogeneity in the data and suggest an improvement in the model fit compared with a standard multinomial logit model.

Table 3. Selection criteria for latent class models.

ID	Models	Log Likelihood	AIC [a]	BIC [b]	Pseudo R^2 [c]
1	Model with 2 latent	−3579	7203	7359	0.11
2	Model with 3 latent	−3549	7168	7405	0.13
3	Model with 3 latent, considering driver seatbelt	−3549	7173	7424	0.11
4	Model with 3 latent, considering day of a week	−3548	7170	7420	0.12
5	Model with 3 latent, type of vehicle SUV	−3547	7167	7417	0.13
6	Standard multinomial logit model	−3674	7371	7446	0.095

[a] Akaike information criterion = $-2(LL - K)$; [b] Bayesian information criterion = $-LL + [(k/2)Ln(N)]$; [c] Pseudo $R^2 = 1 - (LL/LL(0))$.

Based on the information in Table 3, the best AIC value is given by the three-segment model with type of vehicle for class allocation, while the best BIC was given by the model with 2 components. As both models contain a membership class with significant coefficients, these models would be used for comparison purposes. Comparison across models ID 1 and 5 in Table 3 show that model 5 would be chosen as a best fit model due to a higher overall significance of the coefficients for both membership and utility models. Model 5 is superior to model 1 in terms of Pseudo R^2 as well. It is worth mentioning that the simple MNL has the lowest goodness of fit in terms of BIC and Pseudo R^2.

4.2. Distribution of Observations across Various Segments

Additional investigation was conducted to check if the components were characterized by any special response categories. The decision makers select an alternative that produces the greatest utility U, and an individual n chooses an alternative i $U_{in} > U_{jn}$.

The results indicated that although latent 3, for instance, displays a propensity of not buckling up, there is no meaningful evidence that those individuals belong to a specific response category. The share of the second latent is calculated based on the below equation:

$$Share\ of\ segment_{s2} = \frac{exp(class_{s2})}{1 + exp(class_{s2}) + exp(class_{s3})} \tag{4}$$

Again, the value of 1 in Equation (4) resulted from exp (0) for the first component. A statistical summary of the number of observations for each class is presented in Table 4. The highest number of passengers' observations belong to class 3.

Table 4. Individual shares for each segment.

Latent	Number (%)
1	2596 (40%)
2	844(13%)
3	3050 (47%)

5. Concluding Remarks

Seatbelts are one of the most effective ways of saving lives and reducing traffic injuries. However, the presence of non-seatbelt use is still high among vehicles occupants, especially passengers. This study thus sought to unlock the factors influencing passengers' attitudes toward seatbelt usage and to gain insight into the impacts of drivers' seatbelt status on that of the passengers.

Data were collected from various locations in different counties in Wyoming. The seatbelt status of drivers and passengers, along with their demographic information were observed and recorded. The data was filtered to include only those vehicles that had a single front-seat passenger. The response categories of the passenger seatbelt usage were binary: belted versus non-belted passengers.

The results of the study provide a richer interpretation and estimations of the models' parameters. The results of the segment models' parameters suggest that the heterogeneity was related to almost all predictors. Various numbers of latent components, along with various socio-demographic characteristics for class allocations were considered and compared. To come up with the optimal number of segments and class allocation, AIC and BIC were used. The results indicated that the three-component model, with class allocation based on vehicle type resulted in a better fit compared with two components and a constant value for a class allocation.

The heterogeneities were observed by differences across the point estimated and significant levels across the components. Accounting for heterogeneity is especially important as seatbelt usage has been linked to a habitual behavior rather than a conscious choice [20]. As a result, passengers follow the rules developed based on experience, rather than comparing

the benefits and costs of buckling up. Thus, they might be influenced by seatbelt usage of drivers and their gender based on some unobserved characteristics. Thus, it is expected to see heterogeneity in seatbelt usage due to the randomness of passengers' habitual behavior. In addition, we incorporate the vehicle type and license registrations as they are expected to provide information about drivers' and passengers' attributes.

Although the results are, for most parts, in line with past studies, most of those studies ignore the possible unobserved heterogeneities across the observations for seatbelt studies. Not accounting for unobserved heterogeneity could result in possible bias and erroneous results: a purely deterministic model could result in biased estimates and a lack of insight into the true nature of the parameters' effects. This was observed by differences across signs, magnitudes, and significance of the considered components. For instance, we found that the parameters have various significance (e.g., for belted condition of drivers) or opposite signs (e.g., for time of a day) across various components.

In summary, it was found that when drivers are coming from out of the state, traveling a longer distance, the passengers are less likely to be buckled. Moreover, getting a ride as a passenger during weekdays decreases the likelihood of being buckled. Another important finding was the positive correlation between the belting status of drivers and passengers. That highlights the important implication of drivers' behaviors and actions on passengers' attitudes.

Despite the consideration of all plausible drivers and passengers' characteristics, many expected important factors were not considered in this study. Those include, for instance, age, career, and annual income, among other characteristics of drivers and passengers. However, we accounted for those unseen factors through other confounding factors, such as vehicle types or time of traveling. However, future studies are recommended to consider those factors during data collection to provide a better understanding regarding factors affecting seatbelt use.

Study Implication and Recommendations

The findings of this study provide important information for public health organizations to increase seatbelt usage, especially for vehicle passengers. The results indicated that while investigating seatbelt usage, other factors such as driver characteristics should be taken into account.

The psychological elements of a passenger's choice to buckle up include driver characteristics (e.g., passengers are more likely to buckle up when the driver is buckled) and driver gender (e.g., while the gender of the passenger did not have a significant impact on their choice to buckle up, the gender of the driver was found to be important in the passenger's decision). The discussed results need more investigation to identify the underlying causes of the effects. This might be attributed to the level of passengers' trust or lack thereof while riding in a vehicle. For instance, the trust could result from the belting status or gender of the driver, which might contribute to the passenger's choice to buckle up.

More studies are required, especially to study the psychological impacts of the drivers on passengers. We found that time of travel or types of vehicle ownership are other factors contributing to the seatbelt status of passengers. Those factors are likely to be linked to various socio-demographic or socio-economic attributes of passengers. More educational efforts and studies are needed to unlock the real relationship between underlying factors and the likelihood of passengers' seatbelt use.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. NHTSA. Seat Belts. Available online: https://www.nhtsa.gov/risky-driving/seat-belts (accessed on 1 January 2020).
2. Høye, A. How would increasing seat belt use affect the number of killed or seriously injured light vehicle occupants? *Accid. Anal. Prev.* **2016**, *88*, 175–186. [CrossRef]
3. Beck, L.F.; Kresnow, M.-J.; Bergen, G. Belief about seat belt use and seat belt wearing behavior among front and rear seat passengers in the United States. *J. Saf. Res.* **2019**, *68*, 81–88. [CrossRef]
4. Jermakian, J.S.; Weast, R.A. Passenger use of and attitudes toward rear seat belts. *J. Saf. Res.* **2018**, *64*, 113–119. [CrossRef] [PubMed]
5. Briggs, N.C.; Lambert, E.W.; Goldzweig, I.A.; Levine, R.S.; Warren, R.C. Driver and passenger seatbelt use among US high school students. *Am. J. Prev. Med.* **2008**, *35*, 224–229. [CrossRef] [PubMed]
6. Williams, A.F.; McCartt, A.T.; Geary, L. Seatbelt use by high school students. *Inj. Prev.* **2003**, *9*, 25–28. [CrossRef] [PubMed]
7. Aidoo, E.N.; Amoh-Gyimah, R.; Ackaah, W. The relationship between driver and passenger's seatbelt use: A bivariate probit analysis. *Int. J. Inj. Control. Saf. Promot.* **2021**, *28*, 179–184. [CrossRef] [PubMed]
8. Goetzke, F.; Islam, S. Determinants of seat belt use: A regression analysis with FARS data corrected for self-selection. *J. Saf. Res.* **2015**, *55*, 7–12. [CrossRef]
9. Han, G.-M. Non-seatbelt use and associated factors among passengers. *Int. J. Inj. Control. Saf. Promot.* **2017**, *24*, 251–255. [CrossRef] [PubMed]
10. Buchman, T. Does driver seatbelt use increase usage among front seat passengers? An exploratory analysis. *J. Saf. Res.* **2021**, *78*, 170–179. [CrossRef] [PubMed]
11. Train, K.E. Recreation Demand Models with Taste Differences over People. *Land Econ.* **1998**, *74*, 230–239. [CrossRef]
12. McFadden, D. *Conditional Logit Analysis of Qualitative Choice Behavior*; Frontiers in econometrics: New York, NY, USA, 1973.
13. Hess, S. Latent class structures: Taste heterogeneity and beyond. In *Handbook of Choice Modelling*; Edward Elgar Publishing: Cheltenham, UK, 2014.
14. Greene, W.H.; Hensher, D.A. A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transp. Res. Part B Methodol.* **2003**, *37*, 681–698. [CrossRef]
15. Kamakura, W.A.; Russell, G.J. A probabilistic choice model for market segmentation and elasticity structure. *J. Mark. Res.* **1989**, *26*, 379–390. [CrossRef]
16. Scarpa, R.; Thiene, M. Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preferences. *Land Econ.* **2005**, *81*, 426–444. [CrossRef]
17. WYDOT. Seat Belt Survey Analysis. 2019. Available online: https://www.dot.state.wy.us/files/live/sites/wydot/files/shared/Highway_Safety/Annual%20Report/2019%20Survey%20Analysis%20of%20Seat%20Belt%20Use%20in%20Wyoming.pdf (accessed on 1 January 2021).
18. National Highway Traffic Safety Administration. Uniform Criteria for State Observational Surveys of Seat Belt Use. *Fed. Regist.* **2000**, *63*, 46389–46394. Available online: https://www.federalregister.gov/documents/2012/04/05/2012-8137/uniform-criteria-for-state-observational-surveys-of-seat-belt-use (accessed on 30 September 2021).
19. Louviere, J.J.; Hensher, D.A.; Swait, J.D. *Stated Choice Methods: Analysis and Applications*; Cambridge University Press: Cambridge, UK, 2000.
20. Calisir, F.; Lehto, M.R. Young drivers' decision making and safety belt use. *Accid. Anal. Prev.* **2002**, *34*, 793–805. [CrossRef]