

Article

Accounting for Attribute Non-Attendance and Common-Metric Aggregation in the Choice of Seat Belt Use, a Latent Class Model with Preference Heterogeneity

Mahdi Rezapour* and Khaled Ksaibati

Wyoming Technology Transfer Center Laramie, Laramie, WY 82071, USA; khaled@uwyo.edu

* Correspondence: mrezapou@uwyo.edu

Abstract: A choice to use a seat belt is largely dependent on the psychology of the vehicles' occupants, and thus those decisions are expected to be characterized by preference heterogeneity. Despite the importance of seat belt use on the safety of the roadways, the majority of existing studies ignored the heterogeneity in the data and used a very standard statistical or descriptive method to identify the factors of using a seatbelt. Application of the right statistical method is of crucial importance to unlock the underlying factors of the choice being made by vehicles' occupants. Thus, this study was conducted to identify the contributory factors to the front-seat passengers' choice of seat belt usage, while accounting for the choice preference heterogeneity. The latent class model has been offered to replace the mixed logit model by replacing a continuous distribution with a discrete one. However, one of the shortcomings of the latent class model is that the homogeneity is assumed across a same class. A further extension is to relax the assumption of homogeneity by allowing some parameters to vary across the same group. The model could still be extended to overlay some attributes by considering attributes non-attendance (ANA), and aggregation of common-metric attributes (ACMA). Thus, this study was conducted to make a comparison across goodness of fit of the discussed models. Beside a comparison based on goodness of fit, the share of individuals in each class was used to see how it changes based on various model specifications. In summary, the results indicated that adding another layer to account for the heterogeneity within the same class of the latent class (LC) model, and accounting for ANA and ACMA would improve the model fit. It has been discussed in the content of the manuscript that accounting for ANA, ACMA and an extra layer of heterogeneity does not just improve the model goodness of fit, but largely impacts the share of class allocation of the models.



Citation: Rezapour, M.; Ksaibati, K. Accounting for Attribute Non-Attendance and Common-Metric Aggregation in the Choice of Seat Belt Use, a Latent Class Model with Preference Heterogeneity. *Algorithms* **2021**, *14*, 84. <https://doi.org/10.3390/a14030084>

Academic Editor: Frank Werner

Received: 14 January 2021

Accepted: 4 March 2021

Published: 6 March 2021

Keywords: attributes non-attendance; common metric attributes aggregation; latent class; mixed-mixed model; seat belt; traffic safety

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Motor vehicles are a leading cause of death among individual aged 1–54 in the U.S. [1]. Despite the progress in terms of education and laws to motivate individuals to buckle up, the U.S. still has one of the highest traffic death rates per 100,000 population among 20 high-income country members [2]. A lack of protection for vehicle occupants is one of the main causes of the high number of deaths on the roadway. In the U.S., more than half of teens (13–19 years), and adults aged 20–44 who die annually were not buckled up at the time of crashes [3]. That is especially important as the likelihood of passengers being buckled up is significantly lower than drivers. For instance, in Wyoming, while more than 80% of drivers are buckled up, only less than 50% of the front-seat passengers are buckled.

A project was funded in the state to identify underlying factors that persuade the vehicles' occupant to wear or not to wear a seat belt. The first step in reaching this goal was to find most accurate statistical method to identify the underlying factors of the choice to buckle up. The right statistical method is important to account for the preference

heterogeneity of the dataset. That is especially important when dealing with a choice of individuals in making a decision as their choices would be impacted based on various socio-demographic, and environmental characteristics. In our other study, we accounted for taste and scale heterogeneity with a similar dataset. We found that the generalized multinomial logit model is not required at the cost of added parameters [4].

The majority of past studies have been focused on accounting for the preference heterogeneity through the mixed logit model by considering continuous distributions. An extension of mixed logit model has been considered by accounting for taste and scale heterogeneities [5,6]. The latent class model could be considered as an alternative to the mixed logit model by replacing the continuous distribution with a discrete one. For the latent class model, membership in a distinct classes would be used to account for preference heterogeneity [7]. The latent class model assumes homogeneous observations within the same class. However, the assumption of homogeneity has been debated in the literature review and discussed that adding another layer could improve the model fit. That could be implemented by adding another layer of continuous distribution to account for possible preference heterogeneity within a same class [8]. This model has been often called the mixed-mixed logit model in the literature review [9].

The mixed-mixed logit model could be further extended to fix attribute processing rules (APR): by allowing attribute non-attendance (ANA), or aggregation of common-metric attributes (ACMA) [10]. The reason for ANA is that some individuals might ignore some attributes while choosing a specific choice [11]. It should be noted that ANA would not only capture the respondents with a zero sensitivity but also respondents with a low sensitivity [12].

On the other hand, some individuals might exhibit common metrics in making decisions which would be referred as ACMA. Accounting for ACMA and ANA are especially important as it is expected some of the front-seat passengers within the same class might ignore some of the attributes due to some personal characteristics or preferences, or some observations assign similar importance to some other attributes within the same class, and ignoring those characteristics might result in degradation of the model fit.

Thus, this study adds another layer (mixed model) with continuous distribution on top of LC model. It then extends the model by considering ANA and ACMA across front-seat passengers to check if accounting for heterogeneity within a class or considering ANA or ACMA could add values, in terms of goodness of fit, to the models. The findings of this study provide evidence whether accounting for another layer of heterogeneity, along with considering ANA or ACMA, are needed to account for the whole story. That is especially important as we used only individual-specific observations in our dataset. This paper is organized as follows. Section 2, the method section, presents the latent class model along with the considered model specification, then the paper concludes with a summary of the findings.

2. Case Study and Data

The data were drawn from a survey conducted in a western state of the U.S. The data collection was done across 17 counties, and 289 locations in Wyoming in 2019. The data collection process conformed to the criteria highlighted for the state observational seatbelt coalition, which was issued in 2011 by the national highway traffic administration (NHTSA) [13].

The data include various environmental, and demographic characteristics, and seat belt status of vehicles' occupants to see how those factors motivate the occupants to buckle up. Originally there was information related to 18,286 vehicles. However, as the objective of this study was to evaluate the choice of front-seat passengers, and not all the drivers had a front-seat passenger on board, the data were filtered to include only the vehicles that had a single passenger on board. That reduced the number of observations to 6533. Recall, the data also include only individual-specific characteristics, and no information was collected regarding the alternative-specific preferences.

In addition to front-seat passengers, there was information regarding the drivers' characteristics which were incorporated in the analysis. That is because a multitude of reasons could be linked to this choice of a seat belt use for front-seat passengers, and those are not necessarily related to passenger-specific characteristics but also drivers' characteristics.

The individual-specific preference of front-seat passengers, and some of the driver-specific characteristics that were found to impact a choice of passengers in buckling up, were included and highlighted in Table 1. An initial analysis of the dataset indicates that while more than 80% of drivers buckle up, the number is reduced to less than 50% for front-seat passengers.

Table 1. Descriptive summary of important predictors.

Attributes	Mean	Variance	Min	Max
Front-seat passenger belt status, belted versus not belted *	0.32	0.217	0	1
Driver belt condition, not belted versus belted *	0.140	0.120	0	1
Sunny weather condition, sunny as 1, versus others *	0.682	0.216	0	1
Number of lanes, 2 lane as versus a single lane *	1.415	0.243	1	2
Vehicle license, Non-Wyoming residence versus others *	1.552	0.247	1	2
Time of observation: 11:30–1:30 versus others *	0.229	0.176	0	1
Time of observation: 1:30–3:30 versus others *	0.176	0.145	0	1

* Reference category as the Min values.

3. Method

Latent class model could be used as an alternative to the mixed logit model by replacing the continuous distribution with a discrete one. The LC model assumes homogeneous preference across the same class. However, it has been argued that homogenous preference might not hold true, as there might still exist heterogeneity within the same class. Thus, another layer on top of the LC model might be needed to account for the heterogeneity. As the implemented model is an extension of the Mixed-Mixed (MM) model, and the latent class method, the following sections would detail first the standard LC model and then it would discuss its extension.

In the latent class model, it is assumed that the population can be grouped into Q finite number of classes. For the LC model, there are various β_i s across various latent class where the individuals are allocated to classes based on some discrete distribution. While the groups or classes are assumed to be homogenous across the individuals in a same class with specific coefficients, they are heterogeneous across the other considered groups.

The probability of a class q could be written as [14]:

$$Prob(class = q) = \pi_q(\theta), q = 1, \dots, Q \quad (1)$$

where θ is a parameter, which would be used for class allocation. θ might be considered as a constant of 1 or other parameters. Where π_q is the probability of an individual belonging to a class q , where

$$\sum_{q=1}^Q \pi_q = 1 \text{ and } 0 \leq \pi_q \leq 1 \quad (2)$$

Based on various β , and based on g , we would have:

$$f(y_i | x_i, class = q) = g(y_i | x_i, \beta_q) \quad (3)$$

The above equation considers that observations within a same class are homogenous. However, more flexibility could be given to the parameters by varying them based on some

continuous distribution. That could be implemented by adding a mixed layer on top of the latent class model. Therefore, the above model could be modified as [14]:

$$f(y_i | x_i, class = q) = g(y_i | x_i, \beta_{i|q}) \quad (4)$$

where $\beta_{i|q}$ would vary based on some continuous distribution, and it can be written based on some random sampling as:

$$\beta_{i|q} = \beta_q + \omega_{i|q} \quad (5)$$

The simulated maximum likelihood would be used for model parameters, $\beta_{i|q}$, estimation. Now the log likelihood of the resultant would be written as:

$$\mathcal{LL} = \sum_{i=1}^N \log \left[\sum_{q=1}^Q \pi_q(\theta) \frac{1}{R} \sum_{r=1}^R \frac{\exp[\sum_{j=1}^J y_{i,j}(\beta_q + \omega_{ir})x_{ij}]}{\sum_{j=1}^J \exp[(\beta_q + \omega_{ir})x_{ij}]} \right] \quad (6)$$

As can be seen from Equation (6), the process of creating a log likelihood (\mathcal{LL}), could be divided into two parts. First, the class allocation part of $\pi_q(\theta)$, and second, the random sampling of the other parts of the model. Now the process could be summarized as follows:

1. The first process would be related to a mixed part of the model:
 - a. ω estimation: assume $R = 10$ for 3 variables and 100 observations: there would be a matrix with 10 columns and 3 rows, and 100 values: the values would be filled by $\sum_{i=1}^n (\beta_{q_i} + \sigma_i \times \omega_i)$, where ω would be estimated based on pseudo-random numbers or Halton sequences, and σ_i , or the SD of random parameters are values that would be estimated by maximum likelihood, and their initial values would be set by the investigator. Additionally, the initial value of beta would be set by investigator $\rightarrow (\beta_q + \omega_i)$
 - b. The multiplication of the above value by the vectors of observed coefficients would be saved as $XR \rightarrow (\beta_q + \omega_i)x_{ij}$
 - c. The resultant would be multiplied by response $\rightarrow \sum_{j=1}^J y_{i,j}(\beta_q + \omega_i)x_{ij}$
 - d. The exponential of the above values in c would be calculated and would be summed up across the number of J or classes $\rightarrow \exp[\sum_{j=1}^J y_{i,j}(\beta_q + \omega_i)x_{ij}]$
 - e. To have a probability based on the Multinomial logit model, the value in d would be divided by value in c. $\rightarrow \frac{\exp[\sum_{j=1}^J y_{i,j}(\beta_q + \omega_i)x_{ij}]}{\sum_{j=1}^J \exp[(\beta_q + \omega_i)x_{ij}]}$
 - f. There are 10 observations (draws), along with Q columns, related to classes. Thus, the means of each class would be estimated by reducing the dimension of the random draws (R): the average of all the draws over each observation.
 - g. Up to the above steps are related to $\rightarrow \frac{1}{R} \sum_{r=1}^R \frac{\exp[\sum_{j=1}^J y_{i,j}(\beta_q + \omega_i)x_{ij}]}{\sum_{j=1}^J \exp[(\beta_q + \omega_i)x_{ij}]}$.
2. For latent class parts, the steps would be taken as follows:
 - a. Create a vector of γ : this constitutes the initial value of a constant or the heterogeneity point related to a covariate T, which the class allocation is based on. For the first class the value would be set as 0 based on the literature review for model identification.
 - b. Getting the exponential of T, times γ , which discussed in the above,
 - c. The above would be transformed into a probability by dividing the values by the sum of all the components or classes.

- d. The above are related to the $\pi_{n,s}$ which is equal to $\frac{e^{\beta_{n|s}x_{nj}}}{\sum_k e^{\beta_{n|s}x_{nk}}}$ part of the above equation. It should be noted that $\pi_{n,s}$ acts as a constraint with $\sum_s \pi_{n,s} = 1$, so the sum of the probability of each observation across classes would be added up to 1.
3. Now the resultants of item 1 and 2 would be multiplied. This would be the resultant of Equation (6).
4. In order to transform the probability of Equation (6) into a likelihood, the sum of the probability of the individuals would be calculated and set as the likelihood.
5. The log of 4 would be set as log likelihood and would be estimated by maximum likelihood.
6. Now to come up with ANA and ACMA, we put constraints on the means and the SDs of the parameters. For coming up with ANA, we impose restrictions on some attributes' means and for ACMA, we constrain two or more parameter means to be equal.
7. Maximum likelihood would be estimated by the finite-difference method and with the help of Hessian and Gradient.

The standard practice in the use of the LC model is that they assume that all available information related to a choice, which is used by the respondent in making a decision. Additionally, they assume that the factors were assigned to be high importance by the respondents. Those mostly have been assumed in the modeling approach while some studies asked their respondents about whether they considered those characteristics or not [14]. It has been argued that their responses would not be reliable [15]. Additionally, the majority of the past studies assume that all the variables have been used by the respondents. However, there is growing evidence that the respondents (choice makers) might use only a subset of the attributed for making a decision. Those scenarios have been referred to as ANA [16]. On the other hand, the respondents might assign similar values to some attributes due to their similarities and the respondent perception about those characteristics. Those features can be referred to the ACMA.

4. Results

The findings are presented in five models (A to E) in Table 2. The first and second models (standard LC and standard MM model) consider the full attributes attendance (FAA).

It should be noted that normal distributions were assumed for all the considered random parameters. The number of classes was selected based on goodness of fit of Akaike information criterion (AIC). For model C, for instance, although fixed parameters are assumed for all the variables, they were allowed for ANA and ACMA. AIC value was used for determining the number of Q, and as a comparison across various models [10]. As can be seen from Table 2, the AIC is lowest for model E (log-likelihood of -3962). The following paragraphs elaborate on the considered models.

While the first and second models, LC and MM, assume the full attributes attendance (FAA), the other model incorporates a combination of FAA, ANA, and ACMA. For instance, the third model sets constraints on variables such as sunny and day of a week for the first class and driver seat belt status vehicle registration for the second class.

A choice of variables to be incorporated in ANA and ACMA were identified by evaluation attributes one at a time for ANA or ACMA. After identifying variables for both attribute processing rules, they were aggregated and considered in the included models. One of the challenges being observed was for the MM model, especially while considering for more than two classes, due to lack of convergence.

Another point worthy of investigation is which variables to include for ACMA. In the literature review, most variables that belong to the similar category were considered to be incorporated in ACMA. Those include variables, for instance, related to cost or time. These variables mainly have consistent impacts (signs) on the response. Although various predictors were considered to be included for ACMA, they did not result in a model

enhancement and were not justifiable to be used due to lack of interpretation, e.g., driver belt status and vehicle license registration.

Two variables related to time (11:30–1:30 and 1:30–3:30) were found to be a better fit in models. However, as can be seen from all the included models in Table 2, these two variables have opposite signs/directions. Solutions could be proposed including changing the category coding of the variables, ignoring the differences across the signs, or constraining based on their opposite signs. Just to highlight the importance of constraining the variables with various signs, model D (for MM) and model E were proposed.

First, it can be observed that a worst fit model could be observed for the LC model, with no extra layer of heterogeneity, and no ANA and ACMA. An improvement could be observed by including another layer with continuous distribution in the MM model (model B), 7959 for LC versus 7954. Moving to the models with attribute processing rules, ANA and ACMA consideration, although a small amount of improvement could be observed for standard latent class versus the latent class with ANA and ACMA (7959 for a standard latent class versus 7958 for the mixed-mixed model), a model is deteriorated moving from a standard MM to MM model with ANA and ACMA (with a wrong ACMA specification). However, as discussed, for the MM model in D, the parameters were assumed to have similar signs. Model E is presented in which the reverse signs were considered while constraining the parameters. That is a best fit model compared with the other considered models. Although the same approach for the latent class model is not presented in Table 2, it was observed that the AIC was improved for the latent class with reverse signs for ACMA (AIC = 7952), compared with the included model (AIC = 7958).

It is also worth looking at the class membership across the considered models. Class memberships are only significant for ACMA with a wrong approach (models C and D). Additionally, it is worth discussing that while the improvement in model fit for models with ACMA and ANA compared with no attributes processing rules was minor, the differences across the class membership share are hugely different. Consider the standard latent class and the MM models. For those models there is almost equal spread of membership across the two classes. However, moving to the models with wrong specification of ACMA, it can be observed that spread is hugely imbalanced. Finally, the difference between a best fit model considering ANA and ACMA, and all other models is very large.

In summary, the results indicated that the right application of ANA and ACMA attribute processing rules result in an overall improvement in the models' fits. Additionally, it was found that although incorporation of an extra layer of random parameters would not result in a significant enhancement of the model, significant changes could be observed for class shares. This result to some extent was confirmed with the previous studies (see [14]). The results highlight that even in light of ANA and ACMA while using the LC model, an extra layer for taste heterogeneity for improving the performance is still required. It is also worth discussing that the MM models, especially considering ANA and ACMA estimation, took hours, and they faced convergence issue many times.

In summary it is worth discussion the results of a best fit model, model E. First of all, although some uncertainties could be observed in the significance of some t-ratios, those were incorporated for few reasons: first although they were not significant in some models they were significant in others so we kept them despite the uncertainty in the models' parameters estimates. Due to considering two classes, it was possible also that they were not significant in one class and they were in another class. Due to the nature of the dataset, it was challenging to consider only variables in the models that mostly being significant across all the models and across all classes.

Table 2. Summary of the models, t-ratio in parentheses.

<i>Latent class, no ANA, no ACMA (A)</i>		
Attributes	FAA 1	FAA 1
Weather: Sunny	−0.22 (−0.703)	0.19 (0.846)
Day of a week	0.21 (0.679)	0.63 (2.14)
Number of lanes	0.47 (1.033)	0.85 (1.43)
Driver belt status	−0.51 (−1.501)	0.13 (0.49)
Vehicle license plate registration	−0.75 (−2.826)	−0.46 (−2.54)
Time of a day (11:30–1:30)	1.60 (1.408)	−2.66 (−0.88)
Time of a day (1:30–3:30)	−1.41 (−0.658)	0.28 (0.526)
Class probability, γ		−0.16 (−0.134)
Classes share	54%	46%
Log-likelihood		−3964
AIC		7959
<i>Mixed-mixed, no ANA, no ACMA (B)</i>		
Attributes	FAA 1	FAA 2
Weather: Sunny	84.79 (1.422)	0.06 (0.301)
Day of a week	−67.71 (−1.33)	−0.85 (−8.168)
Number of lanes	−106.19 (−1.36)	0.44 (2.844)
Driver belt status	−27.23 (−1.33)	0.33 (2.542)
Vehicle license plate registration	42.60 (1.43)	0.51 (3.534)
Time of a day (11:30–1:30)	−31.37 (1.35)	0.29 (2.587)
Time of a day (1:30–3:30)	133.77 (0.301)	−15.46 (−0.0089)
SD. Driver belt status	91.60 (1.30)	−
SD. Vehicle license plate registration	105.1 (1.37)	−
SD. Time of a day (1:30–3:30)	32.2 (1.39)	−
Class probability, γ		−0.089 (−1.050)
Classes share	52%	48%
Log-likelihood		−3958.7
AIC		7954
<i>Latent class, ANA, ACMA (C)</i>		
Attributes	ANA + ACMA	ANA
Weather: Sunny	—	0.033 (4.666)
Day of a week	—	0.366 (4.666)
Number of lanes	−5.483 (−0.429)	−0.103 (−2.238)
Driver belt status	−4.181 (−0.112)	—
Vehicle license plate registration	−1.239 (−0.461)	—
Time of a day (11:30–1:30)	5.202 (0.414)	−0.298 (−7.104)
Time of a day (1:30–3:30)	5.202 (0.408)	—
Class probability, γ		2.186 (3.871)
Classes share	10%	90%
Log-likelihood		−3969
AIC		7958

Table 2. Cont.

<i>Mixed-mixed, ANA, ACMA (with similar signs) (D)</i>		
Attributes	ANA	ANA + ACMA
Weather: Sunny	—	9.549 (0.140)
Day of a week	—	46.419 (0.033)
Number of lanes	−0.133 (−2.789)	−22.760 (−0.032)
Driver belt status	0.479 (4.516)	—
Vehicle license plate registration	−0.566 (−9.547)	—
Time of a day (11:30–1:30)	−0.271 (−3.647)	−23.192 (−0.033)
Time of a day (1:30–3:30)	−0.271 (−2.85)	—
SD. Driver belt status	0.655 (0.817)	
SD. Vehicle license plate registration	0.130 (0.620)	
SD. Time of a day (1:30–3:30)	0.933 (2.058)	
Class probability, γ		−3.542 (−6.791)
Classes share	3%	97%
Log-likelihood		−3966
AIC		7958
<i>Mixed-mixed, ANA, ACMA (with reverse signs) (E)</i>		
Attributes	ANA + ACMA	ANA
Weather: Sunny	—	0.101 (0.58)
Day of a week	—	0.705 (2.57)
Number of lanes	−0.494 (−2.44)	0.345 (1.24)
Driver belt status	0.703 (2.35)	—
Vehicle license plate registration	−0.889 (−3.11)	—
Time of a day (11:30–1:30)	1.442 (1.40)	−6.742 (−0.42)
Time of a day (1:30–3:30)	−1.442 (−1.42)	—
SD. Driver belt status	1.079 (1.42)	
SD. Vehicle license plate registration	0.518 (1.37)	
SD. Time of a day (1:30–3:30)	2.586 (2.04)	
Class probability, γ		−0.598 (−0.83)
Classes share	65%	35%
Log-likelihood		−3962
AIC		7950

5. Conclusions

Although the benefits of seatbelt use in reduction in road fatalities has been proven, a large number of car occupants do not use their seat belts. Additionally, very few studies have been conducted to study the factors of choosing to wear a seat belt. Previous efforts only focused on traditional statistical analysis for identification of factors to seat belt usage. That is despite the fact that human beings vary in their responses to various stimulus, and thus they respond differently to various attributes.

This paper introduces the latent class model, which is an extension of the MNL model, by dividing the dataset into a few homogenous classes. The analysis is further expanded to the MM model by adding another layer on top of the LC model to account for extra heterogeneity that the standard LC model could not account for. We further

expand the MM model to incorporate ACMA and ANA effects. The process helps to know whether an individual ignores or add attributes. This could be done without asking the respondent about what attributes they did or did not consider while making a decision. On the other hand, the ACMA effect is due to the fact that the differences between some attributes are negligible, and an individual aggregates those characteristics and treats them as identical attributes.

In this paper we implement the discussed model specification in the context of a seat belt dataset, where front-seat passengers choose between wearing a seat belt or not buckling up. The results highlight that adding an extra layer of heterogeneity and considering ANA and ACMA result in an improvement in model fit, and it significantly changes the class allocation shares.

The results highlight that setting the wrong attribute processing rule might just seem like a minor deterioration in a model fit, but it would result in severely imbalanced class allocation shares. This is despite the fact that the class probability and parameters seem to be significant for the wrongly defined model, compared with other models. This highlights the importance of the right consideration for ACMA.

Based on the identified results, some of the front-seat passengers, when choosing to wear a seat belt, assigned similar importance to some attributes while ignoring others. These results were obtained through observation instead of questioning those front-seat passengers regarding their perceptions of various attributes. Additionally, it was found that incorporation of the ANA and ACMA attribute processing rules resulted in an overall enhancement of model goodness of fit.

The results of goodness of fit highlight the importance of answering the questions of how the attributes are considered while evaluating alternatives about seat belt use: whether the individuals across specific class ignore or aggregate some attributes. Accounting for the aforementioned limitation would result in a better fit of the model. We explore these questions across two classes while analyzing the choice of seat belt use.

It should be noted that the findings of this study are specific to the dataset being used. Additionally, it is worth mentioning that all the parameters used in this study were individual-specific and no alternative-specific attributes were considered in the current work.

6. Concluding Remarks

While some of the studies implementing latent class assumed a homogenous utility across the identified classes, others argued that these methods cannot account for the whole story. Thus, there is a growing interest in enhancement of the latent class model by accounting for extra heterogeneity by adding another layer with continuous distribution. The model is named the mixed-mixed method. This model could be further advanced by considering ANA and ACMA. The results of this study highlighted a better goodness of fit of the MM model while considering ANA and ACMA.

It might be argued that a decision being dependent on the vehicle license, for instance, is counterintuitive, as it cannot be said that some would ignore the license information and others would not. Therefore, the vehicle license is likely informative because it says something about the driver, his/her familiarity with the road conditions, and the length of the trip. As a result, these are not attributes of the decision but characteristics of the person making the decision. As such, these could not be called as attributes that one can attend to or not. Having said that, due to the nature of the dataset and lack of availability of choice-specific attributes, that seems the only way of implementation, and here we considered them as attributes. Also those consideration highlighted that work in improvement of the model fit. Additionally, it has been disoced in the literature, when one imposes ANA, the latent class model is a great way of testing for it, but they are no longer classes latent but become a representation of a probabilistic decision rule.

The findings of this this study offer an important insight into the underlying factors of the choice to wear a seat belt. In the present work, we highlighted that the respondents

vary across the same class. An improvement in model fit for the implemented method emphasizes the importance of accounting for data heterogeneity while doing any analysis. The results highlight that some of the front-seat passengers consider time of day as unified variables. On the other hand, while some front-seat passengers do not pay attention to sunny weather condition, and days of weeks, other passengers might take them into consideration.

Author Contributions: Supervision, K.K.; Funding acquisition, K.K.; Methodology, M.R. All authors have read and agreed to the published version of the manuscript.

Funding: The project was funded by the WYDOT.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge that this work is part of project funded by the Wyoming Department of Transportation (WYDOT). The subject matter, all figures, tables, and equations not previously copyrighted by outside sources are copyrighted by WYDOT, State of Wyoming and the University of Wyoming. All rights reserved copy righted in 2016.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Centers for Disease Control and Prevention. *Web-Based Injury Statistics Query and Reporting System (WISQARS)*; National Center for Injury Prevention and Control: Atlanta, GA, USA, 2015.
2. Sauber-Schatz, E.K.; Ederer, D.J.; Dellinger, A.M.; Baldwin, G.T. Vital signs: Motor vehicle injury prevention—United States and 19 comparison countries. *Morb. Mortal. Weekly Rep.* **2016**, *65*, 672–677. [[CrossRef](#)] [[PubMed](#)]
3. National Highway Traffic Safety Administration. *Occupant Protection in Passenger Vehicles. (DOT HS 812 374)*; NHTSA: Washington, DC, USA, 2017.
4. Rezapour, M.; Ksaibati, K. Accommodating Taste and Scale Heterogeneity for Front-Seat Passenger' Choice of Seat Belt Usage. *Mathematics* **2021**, *9*, 460. [[CrossRef](#)]
5. Fiebig, D.G.; Keane, M.P.; Louviere, J.; Wasi, N. The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Mark. Sci.* **2010**, *29*, 393–421. [[CrossRef](#)]
6. Greene, W.H.; Hensher, D.A. Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models. *Transportation* **2010**, *37*, 413–428. [[CrossRef](#)]
7. 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](#)]
8. Greene, W.H.; Hensher, D.A. Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. *Appl. Econ.* **2013**, *45*, 1897–1902. [[CrossRef](#)]
9. Wasi, N.; Carson, R.T. The influence of rebate programs on the demand for water heaters: The case of New South Wales. *Energy Econ.* **2013**, *40*, 645–656. [[CrossRef](#)]
10. Hensher, D.A.; Collins, A.T.; Greene, W.H. Accounting for attribute non-attendance and common-metric aggregation in a probabilistic decision process mixed multinomial logit model: A warning on potential confounding. *Transportation* **2013**, *40*, 1003–1020. [[CrossRef](#)]
11. Hensher, D.A.; Rose, J.M. Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. *Transp. Res. Part E Logist. Transp. Rev.* **2009**, *45*, 583–590. [[CrossRef](#)]
12. Hess, S.; Stathopoulos, A.; Campbell, D.; O'Neill, V.; Caussade, S. It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity. *Transportation* **2013**, *40*, 583–607. [[CrossRef](#)]
13. 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/2011/04/01/2011-7632/uniform-criteria-for-state-observational-surveys-of-seat-belt-use> (accessed on 3 May 2020).
14. Puckett, S.M.; Hensher, D.A. The role of attribute processing strategies in estimating the preferences of road freight stakeholders. *Transp. Res. Part E Logist. Transp. Rev.* **2008**, *44*, 379–395. [[CrossRef](#)]
15. Hess, S.; Rose, J.M. A latent class approach to modelling heterogeneous information processing strategies in SP studies. In Proceedings of the Oslo Workshop on Valuation Methods in Transport Planning, Oslo, Norway, 2–5 November 2007.
16. Hensher, D.A.; Greene, W.H. Non-attendance and dual processing of common-metric attributes in choice analysis: A latent class specification. *Empir. Econ.* **2010**, *39*, 413–426. [[CrossRef](#)]