Article

Latent Class Model with Heterogeneous Decision Rule for Identification of Factors to the Choice of Drivers' Seat Belt Use

Mahdi Rezapour * and Khaled Ksaibati

Wyoming Technology Transfer Center, 1000 E University Ave., Laramie, WY 82071, USA; khaled@uwyo.edu
* Correspondence: mrezapou@uwyo.edu

Abstract: The choice of not buckling a seat belt has resulted in a high number of deaths worldwide. Although extensive studies have been done to identify factors of seat belt use, most of those studies have ignored the presence of heterogeneity across vehicle occupants. Not accounting for heterogeneity might result in a bias in model outputs. One of the main approaches to capture random heterogeneity is the employment of the latent class (LC) model by means of a discrete distribution. In a standard LC model, the heterogeneity across observations is considered while assuming the homogeneous utility maximization for decision rules. However, that notion ignores the heterogeneity in the decision rule across individual drivers. In other words, while some drivers make a choice of buckling up with some characteristics, others might ignore those factors while making a choice. Those differences could be accommodated for by allowing class allocation to vary based on various socio-economic characteristics and by constraining some of those rules at zeroes across some of the classes. Thus, in this study, in addition to accounting for heterogeneity across individual drivers, we accounted for heterogeneity in the decision rule by varying the parameters for class allocation. Our results showed that the assignment of various observations to classes is a function of factors such as vehicle type, roadway classification, and vehicle license registration. Additionally, the results showed that a minor consideration of the heterogeneous decision rule resulted in a minor gain in model fits, as well as changes in significance and magnitude of the parameter estimates. All of this was despite the challenges of fully identifying exact attributes for class allocation due to the inclusion of high number of attributes. The findings of this study have important implications for the use of an LC model to account for not only the taste heterogeneity but also heterogeneity across the decision rule to enhance model fit and to expand our understanding about the unbiased point estimates of parameters.

Keywords: latent class; preference heterogeneity; decision rule heterogeneity; traffic safety; seat belt

1. Introduction

More than 37,000 people died in highway crashes in the U.S. in 2017 alone, out of which 47% were not wearing seat belts [1]. Given the importance of seat belt usage in the enhancement of traffic safety, there has been growing interest in understanding the factors that impact the choice of vehicle occupants in wearing seat belts in order to increase seat belt usage. To achieve that, it is important understand the accurate contributory factors to seat belt usage by implementing the right statistical method that could evaluate unbiased estimated parameters in the choice of seat belt usage.

A drivers' choice for the use of seat belt is expected to be characterized by heterogeneity, which means that data heterogeneity is necessary when making unbiased model outputs estimates. Two main approaches have been proposed in the literature to account for data heterogeneity: the latent class (LC) and mixed models. The mixed model assumes continuous distributions for random parameters, while the LC model relaxes that assumption by using discrete distributions. That is especially important because the choice of random parameter distributions has been subjectively made by investigators, and it is usually challenging to come up with a real distribution underlying random parameters.

The LC model stratifies individuals into various classes. Normally, that entails assigning a constant value or a combination of constant and various socio-economic characteristics, while the first class is normalized to zeros for model identification. A limitation of assigning a constant value for class allocation is that individuals' characteristics are invariant across various choices. In other words, the impact of those individuals' characteristics might not be identifiable when using a constant value for class allocation [2]. To address this shortcoming, it is important to consider using various characteristics for the probability of individuals' assignments to various segments. Thus, the choice of choosing an alternative might be based on a set of socio-economic characteristics.

Still, a limitation of the LC model based on socio-economic characteristics is that all behavioral and individual class membership processes are identical and based on specific attributes or a constant value. Though accounting for taste heterogeneity through a standard LC model usually results in an increase in model fit compared to a standard multinomial logit (MNL) model, this model can only account for variation in marginal sensitivities, and there might still be some heterogeneity due to variations in the choice process by individual respondents [3].

The limitation of the discussed LC model based on various socio-economic characteristics is that the model cannot account for the heterogeneous decision rule. For instance, while some drivers might ignore some characteristics while making a choice, others might use those same characteristics. Not accounting for that heterogeneous decision rule is expected to impact the goodness of a models' fit and, consequently, point estimates. That is especially important because what has been identified as a normal taste heterogeneity might, in fact, be due to decision rule heterogeneity [3].

The heterogeneous decision rule or lexicographic rule is linked to the individuals who value some characteristics so highly that they are not willing to make a tradeoff [4]. By lexicographic choices, we mean that individuals choose an alternative based on some socio-economic or demographic characteristics. It has been noted in the literature that those lexicographic rules can also be used for perception formation in addition to alternative evaluations [5]. The following paragraphs outline a few studies conducted by modifying the standard LC model to account for the heterogeneous decision rule.

Lexicographic rules from choice data related to the few manufacturers of televisions were used [4]. The process was implemented by two main steps: first, the results of the finite mixture model were used to identify an initial clustering of observations, and then arbitrary lexicographic rules were used to be assigned to various clusters, and each cluster was assigned to the best of those rules. In another study, the heterogeneous decision rule was considered [3], and a significant gain was observed when moving from the MNL model to the LC model and the LC model which considered the lexicographic rules.

Though many studies related to seat belt use have been conducted in the literature, almost all those studies used traditional techniques that could not account for the heterogeneity in their datasets. A study was conducted to investigate seat-belt-wearing compliance across road users in Malaysia. Gender, time of a day, and type of vehicle were some of the factors found to impact the use of seat belts [6]. The study was cross-sectional, and just a descriptive summary of recorded observations was presented. The impact of demographic factors on seat belt use by injured adults in crashes were reported [7]. The data used in that study were from injured adults admitted into a trauma center. Standard logistic regression was used for the purpose of the analysis. The results indicated that drivers making a greater yearly income, female drivers, and white drivers are more likely to have their seat belts on while involved in crashes.

Occupation, education, driver age, gender, type and make of vehicle, road surface condition, and type of roadway were some of the factors that were found to impact the likelihood of seat belt use in Iran [8]. The collected data were regarding a total of 1427 motor vehicles, and a descriptive summary of the data was presented in the study. Occupant seat belt use in the Ghanaian University campus was recorded [9]. The data were collected via an unobstructive survey by collecting the information from vehicles. It was found that

vehicle type and gender were some of the factors that impact the usage of seat belts. A chi-square test was used to establish the relationship between seat belt use and vehicle type. In another study, the impact of belted drivers on belting status of the passenger was evaluated, and a standard mixed model was used. A total of 33,310 vehicles in Tennessee were described for data collection. The results implied that a seat belt user can be heavily impacted by vehicle type, sex, and driver's belt status [10]. Additionally, two other studies were conducted by the authors of this study that accounted for taste and scale heterogeneity [11] and that accounted for attribute non-attendance and common-metric aggregation [12].

This current study was conducted by extending the standard LC model to an LC model that accounted for the decision rule. The idiosyncratic differences in the choices made by drivers were hypothesized to be linked to various socio-economic characteristics. The consideration of the observations' allocation based on various socio-economic characteristics could provide further insights about the underlying factors that shape the distribution of the parameters.

To have a better understanding of the importance of accounting for the heterogeneous decision rule, a comparison was made between the standard LC model and the models that accounted for that effect. The data section describes the data used in this study, while the method section details the implemented methods. The results and discussion sections describes the obtained results.

2. Data

The dataset was collected in 2019 across 17 counties in 289 locations in Wyoming. 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 seat belt survey issued in 2011 by the National Highway Traffic Safety Administration: the survey followed the uniform criteria for the state observational survey of seat belt use [13].

The data framework used in this study utilized various demographic and environmental characteristics that might motivate drivers to buckle up. There were 18,286 driver observations considered in this study. A summary of the attributes and levels of various contributory factors, along with related descriptive summaries of the important predictors, is presented in Table 1. The data were collected by the observers in two hours span from 7:30 a.m. to 5:30 p.m. (7:30–9:30 and 3:30–5:30). As can be seen from Table 1, while most of the drivers were buckled, a significant proportion of the passengers were unbuckled (mean = 0.21).

Table 1. Summary statistics of the considered predictors.

Attributes	Mean	Min	Max	Variance
Response, non-belted * versus belted	0.21	0	1	0.167
Driver gender, male *, versus female 2	1.32	1	2	0.217
Population, urban * rural	1.74	1	2	0.192
Sunny versus others *	0.71	0	1	0.206
Type of vehicle: Van versus others *	0.29	0	1	0.208
Type of vehicle: SUV versus others *	0.06	0	1	0.208
Type of vehicle: Pickup truck versus others *	0.41	0	1	0.241
Time: 3:30–5:30 versus others *	0.26	0	1	0.191
Time: 1:30–3:30 versus others *	0.15	0	1	0.130
# of lanes 1 * and 2	1.442	1	2	0.247
License plate registration, Wyoming plate as 2 versus others as 1 *	1.43	1	2	0.601
Day of a week, 1 weekend B *, 2 weekdays	0.86	0	1	0.121

* reference category.

Various weather conditions such as clear, foggy, and rain were available for a single predictor. Those predictors were converted into dummy variables, and their significance was checked in the model. Various car types such as van, SUV, and pickup truck were available and considered in the analysis; again, dummy variables of those vehicle types were considered for the statistical analysis assessment. Finally, it should be noted that all the considered variables were individual-specific, and no choice-specific variables were considered in this study.

3. Methodology

This section discusses the implementation of the LC model in detail. As the method could be written as an extension of the multinomial logit model, the description of that model would be presented first. Consider a decision maker making a choice from a finite set of alternatives as J while there are random vectors and utility $(U_1, \ldots, U_j)$; the probability of choosing i from a choice set of M is given by:

$$P_i = \Pr(U_i > U_j), \quad for\ all\ i \neq j \in M \tag{1}$$

where U could be written based on V (systematic utility) and ε (random utility). V is a function of some observable covariate vectors and unknown parameters that can be estimated, while ε contains unobserved determinants of utility, which follows Gumbel distribution type I. The choice model can be solved as:

$$\Pr(y_n = 1) = \frac{1}{1 + e^{-(v_{nA} - v_{nB})}} = \frac{e^{V_{nA}}}{e^{V_{nA}} + e^{V_{nB}}} \tag{2}$$

where the individual n selects a choice across the alternatives of "*A*" as 1, or $y_n = 1$, and "*B*" as 2.

Now, by considering the assumption that the error terms are identically distributed [14] (in addition to the Gamble scale parameter being set as 1 and the probability for utility i being based on the logit transformation ($\frac{1}{1 + e^{-\mu\varepsilon}}$), the MNL model could be written as [15]:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{v_{nj}}} \tag{3}$$

where the above parameters (especially for the LC model) could be defined as follows: J is the total number of alternatives and V_{ni}, which was defined earlier, is a function of the vector of coefficients x_{ni} and coefficient estimates of β.

Thus, a decision maker could choose an alternative *i* with the maximum utility, and if, for instance, $U_{buckled} > U_{unbuckled}$, a driver would choose to be buckled. As the implemented method in this study was an extension of the standard latent class model, the following paragraphs describe the LC model's specifications.

A latent class model can be used to divide heterogeneous data into Q homogenous data segments, where each class has its own parameters. Compared with the mixed model, the LC model is less restrictive because it can identify classes without any predefined assumptions. The parameters for the first class should be normalized to zero to allow for the segments' identification. When allowing for data with Q different elements, the log-likelihood would be written as:

$$LL(y_i | X_i, \theta) = \sum_{N=1}^{n} ln \sum_{q=1}^{Q} \pi_{iq}(\gamma) \prod (P_{ni|q})^{y_{ni}} \tag{4}$$

where π_q is the probability of an individual i belonging to a class q (class allocation probability), $\sum_{q=1}^{Q} \pi_{iq} = 1$, and $0 \leq \pi_{iq} \leq 1$; P_{ni} was given by Equation (3). A few observations can be made from Equation (4). First, the LC model is a probabilistic model

that aggregate the class allocation probability of π_{iq}, and individual probability of $P_{ni|q}$, where $P_{ni|q}$ is based on the MNL model.

Now π_{iq}, which is the multinomial logit model, would be written as [16]:

$$\pi_{iq}(\gamma) = \frac{exp(z_i^{\mathsf{T}}\gamma_q)}{\sum_{q=1}^{Q} exp(z_i^{\mathsf{T}}\gamma_q)}, \gamma_1 = 0 \tag{5}$$

where z_i denotes a vector of socio-economic characteristics determining the class assignment and γ_q represents associated parameters that are a set of estimated constant and various coefficients that can be used for computing class probabilities. Normally, for a standard LC model, similar socio-economic and constant values would be used for all class allocations.

However, due to various behavioral process, various classes might be dependent on various attributes, so each class might need to be identified based on various socio-economic characteristics, and considering all parameters equally for class allocation could result in unnecessary model complexity and biased model parameters estimates. After employing the aforementioned techniques, the goodness of fit of the Bayesian information criterion (BIC) could be used for the evaluation of a model's fit.

Lexicographic choices are defined as a set of choices that allow a respondent to choose an alternative that is superior with particular attributes [17]. For instance, a lexicographic analysis could show whether a respondent chose one option over another while considering similar characteristics. Various reasons have been assigned to lexicographic choices, including the simplification of the choice task and a real preference [17]. However, as the simplification of the choice was not practical for our case study due to limited numbers of alternatives, it can be said that the lexicographic choice was a result of a real preference.

Finally, the parameter estimations are based on the unconditional probabilities for the LC model by plugging Equations (3) and (5) into Equation (4), which is written as follows:

$$P_i = \sum_{q}^{Q} \pi_{iq} \prod_{j}^{J} [\frac{exp(x_{ij}^{\mathsf{T}}\beta_q)}{\sum_{j=1}^{J} exp(x_{ij}^{\mathsf{T}}\beta_q)}]^{y_{ij}} \tag{6}$$

As the likelihood of the above equation does have a close-form solution, no simulation like the mixed model is needed. In an analysis, for π_{iq}, instead of having a similar characteristic across various classes in the Equation (4), the values vary across the segments.

Model Parameter Estimations

The description of model parameter estimates is based on Equation (6) and could be summarized in two parts: estimating π_{iq} (class allocation probability) and based on Equation (5), and $(P_{ni})^{Y_{NI}}$ in Equation (3). The parameter estimates based on the above discussion can be highlighted in a simple form as follows:

1. The description below is related to (π_{iq}) in Equation (6), which is described in detail in Equation (5). This parameter is based on z_i^{T} and γ_q.

    a. Initial values of γ_q, which constitute initial values of a constant and the heterogeneity point estimates related to initial values of covariate for each class would be created. The accommodated matrix can be saved and called T.

    b. Now, z_i^{T} (related vectors of γ_q) times γ_q would be saved as G. The exponential of G is made to be $exp(z_i^{\mathsf{T}}\gamma_q)$. For our case, the values were related to various socio-economic factors and a constant value of 1 for the intercept. As discussed, the value for the first group for all socio-economic factors and the intercept was fixed as 0 for parameter estimations.

    c. The number in b is transformed into a probability by dividing the value of b by the sum of all the segments: $\frac{exp(z_i^{\mathsf{T}}\gamma_q)}{\sum_{q=1}^{Q} exp(z_i^{\mathsf{T}}\gamma_q)}$.

2. The following description is related to $\prod (P_{ni|q})^{y_{ni}}$ in Equations (4) or (6):

    a. Q is number of columns, which is equal to number of classes that are created.

    b. The initial values of model parameters (β_q) can be multiplied by vectors of the related parameters (x_{ij}^T). The initial values might be estimated based on a normal multinomial logit model. The result is $exp(x_{ij}^\mathsf{T}\beta_q)$.

    c. The result is divided by the sum of all the segments as probability $\frac{exp\left(x_{ij}^\mathsf{T}\beta_q\right)}{\sum_{j=1}^{J} exp\left(x_{ij}^\mathsf{T}\beta_q\right)}$.

    d. The multiplication of the result of 2.c times the response is stored as pp (log power rule). This is the second part of Equation (4): $(P_{ni|s})^{Y_{NI}}$.

3. The result of 2.d times the result in 1.c is stored as log likelihood; see Equation (6).

4. The result in 3 would be maximized by finding the optimum values through Hessian and gradient models with the help of finite-difference methods like the maximum likelihood method. In the standard LC method, all classes are modified based on the discussed π_{iq}. However for the LC model that considers the decision rule, the easiest way is to set a constraint on some variables within a class at zeroes so that their class allocations are not dependent on some socio-economic characteristics.

4. Results

It has been documented in the literature that the choice number of latent classes must be selected by the researchers [18]. Thus, we tried various numbers of classes with different variables for class assignment. The number of classes and variables to be considered for class assignments were chosen based on the goodness of fit and interpretability. In addition to the constant value, the three attributes of road classification, vehicle plate registration, and vehicle type of pick up were considered for class allocation.

Three models were considered for a comparison. The first model was the standard MNL model with no random heterogeneity in sensitivities. The second model was the LC model including three classes with no treatment for class assignment. Finally, the third model was treated by a single lexicography rule for only the second class. Decision rule treatment is important when employing an LC model, as otherwise the question of whether the decision makers made their choices based on all the attributes described in the model or some individuals used only some of those criteria would be raised.

A few observations are worthy of discussion. First, an improvement in the model fit could be observed when moving from the standard MNL model (BIC = 18,155) to the standard LC model (BIC = 18,145) and moving from the standard LC model to the LC model with a decision rule adjustment (BIC = 18,134). Additionally, it was clear that with the removal of a parameter in the LC model with the decision rule, the log likelihood decreased by one point, highlighting an improvement in model fit. The classes with the highest and lowest numbers of observations for the second and third models were related to classes 3 and 1, respectively (# of samples).

After checking for the similarity of the second class of the second model, it was observed that only 8 out of 242 observations of the second class of the third model is like the second class of the second model. Thus, after constraining pickup truck for the second class of the third model, allocated observations for all classes were distorted. That could also be observed via changes in assignments and magnitudes of the class probability for license plate, for instance. Because the best model in terms of the goodness of fit was the third model, we now further discuss that model's results.

As can be seen from Table 2, for respondents who did not use Lex.class2_pickuptruck in the choice task (second class of the third model), the pickup type of vehicle had a negative impact on their choice of being unbuckled ($\hat{\beta} = -1.43$); meanwhile, for the drivers who used Lex.class3_pickuptruck (class 3) as their class assignment, a pickup truck had a higher valuation in their choice of being unbuckled ($\hat{\beta} = 0.73$).

However, the results for the non-lexicographic group (first group) highlighted that the impact of the pickup truck was not different that 0. Additionally, time of a day of 3:30–5:30

had a lower valuation of being unbuckled for those drivers who used the lexicography of the pickup truck for their choice ($\hat{\beta} = 0.8$) compared with those drivers who did not ($\hat{\beta} = 4.27$).

Table 2. The results of the considered models. MNL: multinomial logit; LC: latent class; BIC: Bayesian information criterion.

	1st Model: MNL			2nd Model: Standard LC Model			3rd Model: LC Model Considering the Decision Rules		
	Estimate	Std. Error	Pr(>|z|)	Estimate	Std. Error	Pr(>|z|)	Estimate	Std. Error	Pr(>|z|)
Class I									
Driver gender	−0.28	0.039	<0.005	−0.62	0.143	0.000	0.13	0.504	0.79
Roadway classification	−0.79	0.033	<0.005	0.72	0.266	0.007	12.67	6.851	0.06
sunny weather	0.14	0.042	<0.005	0.57	0.163	<0.005	−14.39	7.022	0.04
Type of vehicle: van	−0.26	0.054	<0.005	−0.37	0.161	0.023	−0.43	1.074	0.69
Type of vehicle: SUV	−0.31	0.096	<0.005	−0.64	0.306	0.037	−2.66	1.434	0.06
Type of vehicle: Pickuptruck	0.37	0.046	<0.005	0.81	0.206	<0.005	−1.59	1.119	0.16
Time of a day: 3:30–5:30	0.19	0.044	<0.005	0.31	0.150	0.042	−13.76	7.106	0.05
Time of a day: 1:30–3:30	0.19	0.054	<0.005	−2.74	1.373	0.046	−2.04	0.959	0.03
No of lanes	0.09	0.033	<0.005	1.32	0.228	<0.005	−1.05	0.642	0.10
Vehicle license registration	−0.10	0.027	<0.005	−1.59	0.214	<0.005	5.24	2.204	0.02
Day of a week	0.17	0.053	<0.005	−1.66	0.315	<0.005	−13.54	6.963	0.05
Class II									
Driver gender	——	——	——	−0.29	0.224	0.196	−0.83	0.530	0.12
Roadway classification	——	——	——	−0.06	0.265	0.831	−25.31	23.733	0.29
Sunny weather	——	——	——	−0.25	0.266	0.340	3.29	1.440	0.02
Type of vehicle: Van	——	——	——	−0.38	0.318	0.228	−0.52	0.678	0.44
Type of vehicle: SUV	——	——	——	−0.31	0.563	0.586	0.37	1.060	0.73
Pickuptruck	——	——	——	−0.39	0.354	0.266	−1.43	0.786	0.07
Time of a day: 3:30–5:30	——	——	——	−0.29	0.250	0.249	4.27	1.268	0.00
Time of a day: 1:30–3:30	——	——	——	13.95	68.479	0.839	−0.55	0.759	0.47
No of lanes	——	——	——	−14.05	68.476	0.837	12.01	10.985	0.27
Vehicle license registration	——	——	——	2.04	1.094	0.062	1.97	1.627	0.23
Day of a week	——	——	——	12.87	68.443	0.851	19.94	21.527	0.35
Driver gender	——	——	——	−0.46	0.165	0.005	−0.62	0.112	<0.005
Roadway classification	——	——	——	−4.57	0.751	<0.005	−0.92	0.250	<0.005
sunny weather	——	——	——	0.67	0.210	0.002	1.19	0.301	<0.005
Type of vehicle: van	——	——	——	−0.38	0.180	0.034	−0.46	0.134	<0.005
Type of vehicle: SUV	——	——	——	−0.13	0.269	0.623	−0.34	0.221	0.13
Type of vehicle: pickuptruck	——	——	——	0.04	0.191	0.842	0.73	0.150	<0.005
Time of a day: 3:30–5:30	——	——	——	1.11	0.189	<0.005	0.80	0.142	<0.005
Time of a day: 1:30–3:30	——	——	——	0.17	0.253	0.502	0.52	0.161	<0.005
No of lanes	——	——	——	1.67	0.352	<0.005	−0.10	0.119	0.41
Class III									
Vehicle license registration	——	——	——	0.28	0.052	<0.005	−1.13	0.149	<0.005
Lex. class.3_day	——	——	——	3.52	1.077	0.001	1.10	0.332	<0.005
Lex. class 2_Constant	——	——	——	0.05	0.285	0.864	−3.85	0.570	<0.005
Lex. class 3_Constant	——	——	——	−6.36	1.047	<0.005	−0.60	0.297	0.04

Table 2. *Cont.*

	1st Model: MNL			2nd Model: Standard LC Model			3rd Model: LC Model Considering the Decision Rules		
	Estimate	Std. Error	Pr(>|z|)	Estimate	Std. Error	Pr(>|z|)	Estimate	Std. Error	Pr(>|z|)
Lex. class2_ Roadway classification	—	—	—	0.49	0.138	<0.005	1.88	0.309	<0.005
Lex. class3_ Roadway classification	—	—	—	3.59	0.512	<0.005	1.05	0.129	<0.005
Lex. class2_ vehicle license registration	—	—	—	−1.09	0.220	<0.005	0.62	0.135	<0.005
Lex. class3_ vehicle license registration	—	—	—	0.32	0.079	<0.005	0.47	0.124	<0.005
Lex. Class2_pickuptruck	—	—	—	0.45	0.137	0.001	——	——	——
Lex. class3_pickuptruck	—	—	—	−0.30	0.141	0.032	−0.66	0.128	<0.005
# of samples	—	—	—	3624	2373	12,244	2180	242	15,822
Log Likelihood = −9023; BIC = 18,155				Log Likelihood = −8871; BIC = 18,145			Log Likelihood = −8870; BIC = 18,134		

The results showed that the individuals drivers in the third class, who used a pickup truck to base their choice of being buckled or not while driving, assigned a positive importance to road classification for being buckled ($\hat{\beta} = -0.92$) compared with drivers in the first segment, who made a decision about the choice of seat belt with no rule for class assignment: driving in rural areas increased the likelihood of being unbuckled ($\hat{\beta} = 12.67$) for that segment.

A few observations could also be made regarding the class probability assignment ($\hat{\gamma}$). The probability of being in classes 2 ($\hat{\gamma}_{2.veh\ reg} = 0.62$) and 3 ($\hat{\gamma}_{3.veh\ reg} = 0.47$) increased for those drivers who were driving with a Wyoming vehicle plate registration, with a higher degree for class 2.

Another important finding from the comparison of the two implemented LC models was the change in terms of the significance and magnitude of the parameter estimates, which was also highlighted in the goodness of fit measures. For instance, while the sunny condition increased the chance of being buckled up for the standard LC model, the assignment was reversed for the model with the decision rule consideration. Additionally, while only a single variable in the second class of the standard LC model was not different than zero (p-value > 0.05), the number of important predictors increased to 4 with the modified LC model that considered the decision rule for the same class.

The results of the superior third model showed that for almost all cases, there were mixed results for the parameter estimates across various classes. For instance, driving with a Wyoming license plate and the time of drive were some of the factors found to have impacts on various assignments across classes.

On the other hand, female drivers and driving on weekdays significantly decreased the likelihood of being unbuckled, though only for a single class. Although an increased number of lanes was significant for two classes with an increasing impact of being unbuckled, the assignment was reversed, at the 0.1 significance level, for the modified LC model.

In summary, by just accounting for the heterogeneous decision rule for a single class and across a single variable, a significant, although small, improvement in model fit could be observed when moving from the standard LC model to the modified LC model. Additionally, it was found that lexicographic choice had a significant impact on the selection of an alternative.

5. Conclusions

Two main models of mixed and latent class models have been employed in the literature to account for heterogeneity across individuals. A key limitation of the employment of

parametric distributions in a mixed model is the distribution assumption for the random parameters, which is often set subjectively. The LC model, as an extension of the MNL model, relaxes those assumptions by considering a discrete distribution. While the LC model has mostly used a constant variable for a class allocation in the literature, other studies have used various socio-economic characteristics for class assignments. However, most of those studies considered similar predictors across all class assignments.

One of the shortcomings of the standard LC model is that it considers similar variables or constant values for class assignment across all observations. However, it has been discussed that heterogeneity is not only across individuals but also the class probability assignments to those individuals. Additionally, the constant class assignment assumption is expected to be unrealistic and to be violated, especially due to variations across individual drivers in the choice of seat belt usage.

Three models were considered and compared in this study: the MNL, the standard LC, and the LC model with the heterogenous decision rule. First, various socio-economic characteristics were considered for class allocation probabilities, and it was found that factors including roadway classification, vehicle plate registration, and type of vehicle significantly impacted the class allocations of the individual's observations.

Second, adjustments across variables for the class probability assignments were considered. The results showed that constraining one variable for a single class assignment resulted in a slight improvement the model fit, and significant changes were observed for the level of significance and assignments of few observations. In general, an improvement in model fit was observed when moving from the MNL model to the standard LC, and modified LC models.

The results showed that not accounting for the heterogeneous decision rule is the source of increased errors and may even result in biased and insignificant point estimates for many parameters. Additionally, the impacts of the variables in the assignments of individual drivers into various classes were discussed in detail.

In summary, the results showed that attributes such as a driver's gender, type of vehicle, time of drive, and road classification are some of the factors that impact the choice of seat belt use for drivers. However, mixed results with various assignments and magnitudes were found across almost all predictors.

To improve seat belt usage and have a better understanding of the accurate contributory factors to it, it is important to implement the right statistical method to estimate unbiased parameters for the choice of using a seat belt. It is especially important for policy makers to have this better understanding so that they could could take appropriate steps by targeting the right groups and conducting the correct educational programs.

Author Contributions: K.K., supervision and funding; M.R., analysis. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: There is no available dataset for this study, unless after a justifiable request.

Conflicts of Interest: There is no conflict of interest across the authors.

References

1. NHTSA. Seat Belts. 2020. Available online: https://www.nhtsa.gov/risky-driving/seat-belts (accessed on 29 July 2020).
2. Boxall, P.C.; Adamowicz, W.L. Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach. *Environ. Resour. Econ.* **2002**, *23*, 421–446. [CrossRef]
3. Hess, A.S. Stathopoulos and A. Daly. Allowing for heterogeneous decision rules in discrete choice models: An approach and four case studies. *Transportation* **2012**, *39*, 565–591. [CrossRef]

4.   Jedidi, K.; Kohli, R. Inferring latent class lexicographic rules from choice data. *J. Math. Psychol.* **2008**, *52*, 241–249. [CrossRef]

5.   Bröder, A. Assessing the empirical validity of the "Take-the-best" heuristic as a model of human probabilistic inference. *J. Exp. Psychol.* **2000**, *26*, 1332. [CrossRef]

6.   Fauziana, L.; Atiqah, M.S.; Syukri, Z.A.N. *Identification of Seat Belt Wearing Compliance Associate Factors in Malaysia: Evidence-based Approach*; World Academy of Science, Engineering and Technology: Istanbul, Turkey, 2011; pp. 723–727. [CrossRef]

7.   Lerner, E.B.; Jehle, D.V.; Billittier, A.J.; Moscati, R.M.; Connery, C.M.; Stiller, G. The influence of demographic factors on seatbelt use by adults injured in motor vehicle crashes. *Accid. Anal. Prev.* **2001**, *33*, 659–662. [CrossRef]

8.   Mohammadi, M.; Ansari Moghaddam, A.; Rad, M.; Hashemi Habybabady, R.; Tabasi, M.A. Seatbelt use and related factors among drivers involved in road crashes in Southeast Iran. *Health Scope* **2015**, *4*, e30782. [CrossRef]

9.   Ojo, T.K.; Agyemang, W. Occupants' seatbelt use are related to vehicle type and usage on a Ghanaian university campus. *Int. J. Inj. Control. Saf. Promot.* **2019**, *26*, 145–150. [CrossRef] [PubMed]

10.   Boakye, K.F.; Shults, R.A.; Everett, J.D. Nighttime seat belt use among front seat passengers: Does the driver's belt use matter? *J. Saf. Res.* **2019**, *70*, 13–17. [CrossRef] [PubMed]

11.   Rezapour, M.; Ksaibati, K. Accommodating Taste and Scale Heterogeneity for Front-Seat Passenger' Choice of Seat Belt Usage. *Mathematics* **2021**, *9*, 460. [CrossRef]

12.   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. [CrossRef]

13.   Uniform Criteria for State Observational Surveys of Seat Belt Use. National Highway Traffic Safety Administration Uniform Criteria for State Observational Surveys of Seat Belt Use. *Fed. Reg.* **2000**, *63*, 46389–46394.

14.   McFadden, D. The measurement of urban travel demand. *J. Public Econ.* **1974**, *3*, 303–328. [CrossRef]

15.   McFadden, D. *Conditional Logit Analysis of Qualitative Choice Behavior*; University of California at Berkeley: Berkeley, CA, USA, 1973.

16.   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]

17.   Sælensminde, K. Causes and consequences of lexicographic choices in stated choice studies. *Ecol. Econ.* **2006**, *59*, 331–340. [CrossRef]

18.   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]