Article

Accommodating Taste and Scale Heterogeneity for Front-Seat Passenger' Choice of Seat Belt Usage

Mahdi Rezapour * and Khaled Ksaibati

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

* Correspondence: mrezapou@uwyo.edu

Abstract: There is growing interest in implementation of the mixed model to account for heterogeneity across population observations. However, it has been argued that the assumption of independent and identically distributed (i.i.d) error terms might not be realistic, and for some observations the scale of the error is greater than others. Consequently, that might result in the error terms' scale to be varied across those observations. As the standard mixed model could not account for the aforementioned attribute of the observations, extended model, allowing for scale heterogeneity, has been proposed to relax the equal error terms across observations. Thus, in this study we extended the mixed model to the model with heterogeneity in scale, or generalized multinomial logit model (GMNL), to see if accounting for the scale heterogeneity, by adding more flexibility to the distribution, would result in an improvement in the model fit. The study used the choice data related to wearing seat belt across front-seat passengers in Wyoming, with all attributes being individual-specific. The results highlighted that although the effect of the scale parameter was significant, the scale effect was trivial, and accounting for the effect at the cost of added parameters would result in a loss of model fit compared with the standard mixed model. Besides considering the standard mixed and the GMNL, the models with correlated random parameters were considered. The results highlighted that despite having significant correlation across the majority of the random parameters, the goodness of fits favors more parsimonious models with no correlation. The results of this study are specific to the dataset used in this study, and due to the possible fact that the heterogeneity in observations related to the front-seat passengers seat belt use might not be extreme, and do not require extra layer to account for the scale heterogeneity, or accounting for the scale heterogeneity at the cost of added parameters might not be required. Extensive discussion has been made in the content of this paper about the model parameters' estimations and the mathematical formulation of the methods.

Citation: Rezapour, M.; Ksaibati, K. Accommodating Taste and Scale Heterogeneity for Front-Seat Passenger' Choice of Seat Belt Usage. *Mathematics* **2021**, *9*, 460. <https://doi.org/10.3390/math9050460>

Academic Editor: Armin Fügenschuh

Received: 1 February 2021

Accepted: 20 February 2021

Published: 24 February 2021

Keywords: choice modeling; mixed model; scale heterogeneity; taste heterogeneity; generalized multinomial logit model; front-seat passengers

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

More than 37,000 people died in highway crashes in the U.S. in 2017 alone, of which 47% were not wearing seat belt [1]. Seat belt usage could reduce both fatal and non-fatal injuries by 60% among front-seat occupants, and 44% for rear seat occupants [2]. Despite efforts regarding the importance of seat belt usage, a significant portion of vehicle users do not buckle up. The number of passengers that do not wear seat belts is much higher compared to the drivers. For instance, in Wyoming, more than 80% of drivers buckle up but less than 50% of the passengers used seat belt while getting ride.

It is known that the individuals would base their choices about wearing seat belt based on various attributes, which might show itself with substantial heterogeneity. The main drawback of not accounting for that heterogeneity is linked to biased, and often erroneous estimates of the attributes, and consequently lack of understanding of underlying effects of the point estimates.

Traditionally, the multinomial logit model (MNL) has been employed to identify factors to the choices of individuals. However, the main drawback of this method, in simple words, is that the point estimates are based on the average of all observed preference regarding a specific attribute. That idea ignores the fact that while most of the observers favor preference based on an attribute's characteristics, a significant subpopulation might favor other aspects of an attribute. That approach would miss the real performance of individuals, and consequently would lead to unrealistic point estimates of the parameters. That is especially important while examining the traffic safety study, such as factors to the choice of seat belt use as the primary objective would be to identify unbiased point estimates with an objective of addressing the factors to improve the roadway safety [3].

Mixed model, as extension of the MNL model, has gained much popularity by allowing the observed attributes coefficients to vary based on predefined distributions across observations [4]. However, the standard mixed model assumes the idiosyncratic error term is still independent and identically distributed (i.i.d). Various approaches have been implemented to address the shortcoming of the standard mixed model. A popular technique is to change the means of the random parameters based on some observed attributes. The implementation is expected to result in a significant improvement in the model's fit. However, choice modelers still favor accounting for heterogeneity using unobserved heterogeneity [5]. It was discussed that in the presence of a non-trivial amount of scale heterogeneity, the implemented model without accounting for the scale might overstate the sensitivity degree of individuals heterogeneity [6].

In this study, the MNL was extended to the mixed model to account for heterogeneity in taste and scale. The MNL, with its extensions, are employed on the choice data of seat belt use related to the front-seat passengers. An improvement in the model fit due to the implementation of the generalized MNL (GMNL) is expected when there is a possible presence of the extreme individuals exhibiting nearly lexicographic preferences or those showing very random behaviors.

It should also be noted that it has been objected that the variation in absolute sensitivities (scale heterogeneity) could not be separated by heterogeneity in individual coefficients, or taste heterogeneity, and the improved fit due to accounting for the so-called scale heterogeneity is only due to more flexible distributions, rather than the scale heterogeneity itself [6]; in other words, by treating the scale, we would indeed treat the taste intensities. That idea is in line with earlier work stating that the scale and taste heterogeneity are confounded and change in one would result in change in the other factor [7].

Also, it was discussed in the literature that accounting for scale heterogeneity in the absence of preference heterogeneity only provide limited model fit improvement, while accommodating both scale and taste heterogeneity result in statistically superior model [3]. However, all the discussed points might be specific to the type of dataset being used for the analysis, and amount of randomness and extreme choice being recorded.

The following paragraph would outline some of the recent study that implemented the GMNL in various studies. The GMNL was used to model air travel choice [8]. The results highlighted that the GMNL would outperform the mixed logit (MIXL) model and accounting for both scale and taste heterogeneity help to identify the unique consumer segment. The study tested how providing information regarding the ecological processes impact the public preferences [9]. The GMNL was used and it provided information on scale and unobservable heterogeneity.

The current study is conducted to estimate the parameters to seat belt use in a more reliable way. As a result, we would evaluate in this study, if accounting for heterogeneity in scale, in addition to taste, is needed. Moreover, the focus of this study is in methodological approach of the implemented techniques so a significant proportion of this study would dedicate to that matter.

The remainder of this study is structured as follows: the data section would present the data used in this study, while the method section would discuss the implementation

of the MNL and its extension to the mixed logit model (MIXL), scale multinomial logit (SMNL) and GMNL model. The final two sections will cover the results and conclusions.

2. Data

The dataset collection was performed in 2019 across 17 counties, at 289 locations, in Wyoming. The observers who collected the data were trained in a classroom before any data collection. That 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, 23 CFR [10].

Originally, there were 18,286 vehicles' observations. As the objective of this study was to only evaluate the front-seat passengers' seat belt 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 seat belt status, the total number of dataset observations was reduced to 6533. In the dataset, there were attributes regarding drivers in addition to passengers.

In the analyses, various drivers' and passengers' characteristics were considered. The consideration of both factors would help to unlock important information regarding the interactions of drivers and passengers with regard to the passengers' seat belt use. A summary of the attributes and levels of various important contributory factors, along with related descriptive summaries 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' seat belt statuses. Driver gender was included in Table 1, as it was found that this predictor is important in predicting the passengers' seat belt status. For instance, 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 *	0.679	0.218	0	1
vehicle characteristics				
pickup versus other *	0.329	0.221	0	1
SUV versus others *	0.073	0.068	0	1
Van versus others *				
Vehicle license, Non-Wyoming residence versus Wyoming resident *	0.60	0.600	0	1
Environmental characteristics				
Sunny versus other *	0.681	0.217	0	1
Day of a week, weekdays versus weekend *	0.799	0.161	0	1
Driver characteristics				
Driver belt condition, belted versus not belted *	0.141	0.121	0	1
Driver gender, female versus male *	0.260	0.192	0	1
Environmental condition				
7:30–9:30 versus other *	0.158	0.133	0	1
11:30–1:30 versus other *	0.23	0.176	0	1

* Reference category as 0.

3. Methods

The mixed logit (MIXL) model, as an extension of the MNL, allows for random coefficients on observed attribute, but still considers the i.i.d error terms. However, it has been argued that for some observations the scale of the error is greater than others [11], which might invalidate the i.i.d error term assumptions. The GMNL was described as a model addressing the MIXL model shortcoming by accounting for lexicographic preference of extreme individuals. That would be accounted for by giving more flexibility to the distribution in case of extreme values.

Here the interest for random parameters analysis or MIXL model is to give flexibility to the parameters to vary across population based on some pre-specified distributions. The necessity of the parameters to vary across observations would be evaluated by the significance of the standard deviation of the random parameters.

MIXL generalizes the MNL model, while the generalized multinomial logit (GMNL) model would nest the scale MNL (SMNL), and MIXL models. As the GMNL model is an extension of the MNL model, this section will first discuss the general form of MNL, and then it will extend it to the implemented models of MIXL, GMNL, and SMNL. The random utility for a person i , and an alternative j would be written as:

$$U_{ij} = X_{ij}^T \beta_i + \varepsilon_{ij}, \quad i = 1, \dots, N, \quad j = 1, \dots, J \quad (1)$$

where x_{ij} is an $i \times j$ matrix of observed attributes, ε_{ij} is the idiosyncratic error term being i.i.d or extreme value type I. In case of the MIXL model and if β_i is random, the parameters follow the continuous density $f(\beta_i|\theta)$, where θ are the parameters of the random parameters. In summary, the utility could be divided into three components: deterministic, random and error term components.

The choice probability for Equation (1) is:

$$P(j|X_n) = \frac{\exp(\beta x_{nj})}{\sum_{j=1}^K (\beta x_{nj})} \quad (2)$$

Now in case of β_i being random, β_i would follow distributions such as MVN (β, σ), and could be defined as follows:

$$\beta_i = \beta + L\omega_i \quad (3)$$

where ω_i are draws from a predefined distribution such as $N(0, I)$, variance of β_i is σ which could be separated to its Cholesky factors of L (lower triangular Cholesky factor of σ) and its transpose, as L^T so $\sigma = L L^T$. L would be a lower triangular considering the covariance matrix of the random parameters.

In case of correlated random parameters, the Cholesky matrix based on the covariance matrix, would accommodate the correlation values between two random parameters in the lower Cholesky matrix, off-diagonal values. However, for a normal situation with no correlation across random parameters, the matrix L is only the diagonal matrix of standard deviations.

Now by considering random parameters, the Equation (1) would be rewritten as: $U_{ij} = X_{ij}^T (\beta + L\omega_i) + \varepsilon_{ij}$, and consequently the probability in Equation (2) could be written as.

$$P(j|X_n) = 1/D \left(\frac{\exp\left(X_{ij}^T (\beta + L\omega_i)\right)}{\sum_{j=1}^K \left(X_{ij}^T (\beta + L\omega_i)\right)} \right) \quad (4)$$

where D is the number of random draws from a predefined distribution.

The GMNL, as a general form of the MIXL, could be written as [5]:

$$\beta_i = \sigma_i \beta + [\gamma + \sigma_i(1 - \gamma)] L \omega_i \quad (5)$$

where σ_i is individual-specific scale of the error term, γ is a scalar value controlling how the residual taste heterogeneity of $L\omega_i$ varies with scale. Although primarily $\gamma \in [0, 1]$ by using logistic transformation, it was found later that even $\gamma < 0$ or $\gamma > 1$ would cause no problem in behavioral interpretation [12]. As can be seen from Equation (5), γ controls how the residual taste heterogeneity variance (η) varies with scale heterogeneity of σ .

The above could be turned into the GMNL-I ($\beta_i = \sigma_i\beta + L\omega_n$), or G-MNL-II ($\beta_i = \sigma_i(\beta + L\omega_n)$) for $\gamma = 1$, and $\gamma = 0$, respectively [5].

It should be noted from Equation (5) that the MIXL is a general case of the GMNL-II, where $\sigma_i = 1$ so consequently $\beta_i = \beta + L\omega_i$. On the other hand, the SMNL is a special case of GMNL where variance of ω is zero and thus β_i is:

$$\beta_i = \sigma_i\beta + \varepsilon_i \quad (6)$$

where ε_i is residual heterogeneity, which would not be explained by scale heterogeneity [5]. In the above formula we have σ_i instead of σ , which is due to the fact that the scale parameter would vary across various parameters due to scale heterogeneity consideration. The front-seat passenger utility consequently, based on the SMNL, could be written as: $U_{nj} = (\beta\sigma_i)X_{nj} + \varepsilon_{ij}$.

As can be seen from Equation (5), the parameters' estimates are proportional across individual front-seat passengers by the scaling factor of σ_i . To constrain σ_i to be positive, the log-normal distributions with mean of $\bar{\sigma}$ and standard deviation of τ would be set, so σ_i would be written as:

$$\sigma_i = \exp(\bar{\sigma} + \tau\varepsilon_i) \quad (7)$$

It has been noted that the performance of the above algorithm would be improved by setting $\bar{\sigma}$ as follows [5]:

$$\bar{\sigma} = -\log\left[\frac{\sum_{n=1}^N \exp(\tau\varepsilon)}{N}\right] \quad (8)$$

Also in Equation (7), various front-seat passengers, or intercept could be incorporated in σ_i so we would have:

$$\sigma_i = \exp(\bar{\sigma} + \delta z_i + \tau\varepsilon_i) \quad (9)$$

where δ is the estimated vector of attributes related to various individual characteristics of z_i .

Now the likelihood would be estimated by summing up the probability by nesting the MIXL and the SMNL as:

$$\hat{P}_n = \frac{1}{D} \sum_k^K \prod_j^J \left(\frac{\exp(\sigma_k\beta + [\gamma + \sigma_i(1 - \gamma)]L\omega_k)X_{nj}}{\sum_{j=1}^J \exp(\sigma_k\beta + [\gamma + \sigma_k(1 - \gamma)]L\omega_k)X_{nj}} \right)^{y_{nj}} \quad (10)$$

Model Parameters' Estimation

The process of estimating the model parameters will be presented in three subsections and related parts. The main challenge for model parameter estimation is related to the preparation of the data: the preprocessed steps for data has been referred to long or wide dataset in the literature. That is especially important as all the data information in this study is related to the individual-specific characteristics. However, this study does not detail that process, but only briefly focuses on a general process as follows:

- A. The description below is related to Equation (9).
 - a. Epsilon (ε) would be generated by truncated random parameters, between the range of -2 and $+2$ as the large value of tau would cause numerical issues for draws of epsilon [5].
 - b. $\bar{\sigma}$ would be estimated based on Equation (8). An initial value would be given to tau which would be updated through iterations by the maximum likelihood method in the later process. Tau times ε in (a) and getting the mean of the

- estimations for each row by dividing by the number of observations, and consequently getting the $-\log$.
- c. Vector of intercept or an explanatory variables time initial values of heterogeneity parameters of $\delta, \delta z_n$. For instance, in case of an intercept, vector of 1 for z_n would be multiplied by initial value of δ , or in case of other observed attributes z_n would be vectors of those attributes. Also, if no variable was considered, the value of δz_n would be left as zero.
 - d. Now ε would use part of epsilon for each observation. It would be multiplied by initial value of tau to make $\tau\varepsilon_n$.
 - e. The sum of d, c, and b are related to estimation of σ_n in Equation (9). It should be noted that the exponential transformation was used to constrain the scale parameter of σ_n to be positive [5].
- B. Now β would be created based on Equation (5), for each attribute. The process could be summarized in 3 parts of $\sigma_i\beta$, $\gamma L\omega_i$, and $\sigma_n(1 - \gamma)L\omega_i$.
- a. β times σ_n , which was estimated in part A, as $\sigma_n\beta$
 - b. Initial value of γ times initial value of standard deviation (SD) or L of a variable, times the allocated draws related to that variable as $\gamma \times L\omega_i$. It should be noted that ω_i could be based on Halton draws (prime numbers process), or random parameters based on some preassigned distribution.
 - c. Again, we would use σ_n in A, times $(1 - \gamma)$, times SD times the related draws as $\sigma_n(1 - \gamma) \times L\omega_i$.
 - d. It is worthy to mention that part B has all the required parameters for the output of our models including variables such as β, γ and SD , which could be updated and estimated based on the maximum likelihood algorithm.
 - e. It is also noteworthy that the process in A and B would be employed in a loop for each observation by using their related draws, and while all parameters, e.g., tau, or gamma, are fixed across the observation, the draws or ω_i would vary based on an observation i.
- C. The information below is related to the final process based on Equation (10).
- a. Now the identified initial estimate in B times the related vector of X_{nj} .
 - b. Converting the process into a probability by dividing the values over the sum and multiplying by response(y), based on the log rule.

Now the maximum likelihood algorithm would be employed by iteration to find optimal values based on assigned Gradient, and Hessian.

Recall that the GMNL is a general form of the MIXL and the SMNL models, where in the SMNL point estimates of beta would be estimated by $\beta_i = \sigma_i\beta$, in MIXL that value is $\beta_i = \beta + L\omega_i$. On the other hand, the SMNL has been described as a parsimonious description of the data compared with the MIXL model as $\sigma_n\beta$ is simpler than $\beta + L\omega_i$ in the MIXL [5]. Also, it is clear that the number of parameters in SMNL is much less compared with MIX and GMNL.

Akaike information criterion (AIC), consistent AIC (CAIC), and Bayesian information criterion (BIC) were considered to choose across GMNL, SMNL and MIX models. The AIC measure has the smaller penalty for increased numbers of parameters, so it is expected to work more favorably for the models with higher number of parameters, GMNL and MIX compared with SMNL.

Various measures were proposed in the literature for model performance evaluation. The measures include the AIC, BIC, and CAIC. However, the reliability of those measures for comparison and which method is more reliable for distinguishing a best fit model has gone under question. For instance, the data were simulated based on the true model of SMNL, MIXL and GMNL [5]. The results highlighted that there is no consistent identification of best fit models so the total number of times that each measure correctly highlight right models were considered to be the best measures. Based on the measures' equation,

it is intuitive to think that some measures would pick more parsimonious models and vice versa.

In the above study, it was found that the AIC correctly picked a model in 61/80 cases, while BIC and CAIC picked right models in 68/80 and 66/80 of cases, respectively. In summary, the study recommended the BIC and CAIC based on the higher numbers of picks.

In this study, we consider looking at the significance of the model parameters, and especially the comparison across the SMNL and the MNL models to come up with a more reliable measure. Also, no absolute best or worse methods would be highlighted based on a sole measure of the discussed fits.

4. Results

It is vital to adequately account for the heterogeneity in the dataset to minimize the biased estimates of model parameters. This study was set forward to check the performance of multiple techniques in accounting for the heterogeneity of the dataset, with an objective of estimating the model parameters in a most reliable way. The methods include the standard MNL, and its extension to the SMNL, the MIXL and the GMNL. Also, correlated random parameters of the GMNL and the MIXL were considered for comparison purposes. It has been discussed that the performance of GMNL and MIXL or SMNL are dependent on the complexity of the data [12], suggesting more sophisticated models would not necessarily result in a better fit.

In this study, various observed attributes were considered for shaping the scale heterogeneity. For instance, for GMNL, vehicle plate registration variable resulted in the BIC of 7729, while an attribute of a day of a week for shaping the GMNL resulted in the BIC value of 7639. On the other hand, weather condition of sunny (BIC = 7596), due to its lowest BIC was kept for scale heterogeneity for both GMNL and SMNL models.

A combination of BIC, AIC, and CAIC and also comparison across the MNL and SMNL were used for making a decision regarding the best fit model. The estimation results are presented in Table 2. GMNL nests both mixed and SMNL in a single model to account for both aspects of heterogeneity. Based on the Equation (5), the $L\omega$ and σ are the key for estimating the β . Transformation of GMNL to MIXL would occur when σ moves toward one, and to SMNL when variance of $L\omega_i$ move toward zero (see Equation (5)). As can be seen from Table 2, the scale parameters τ increases moving from GMNL to SMNL as a possible reason that the weight of all heterogeneity would be on the scale heterogeneity, and random parameters plays no role.

A comparison across standard GMNL and MIXL, highlighted that although log likelihood (LL) improves for the GMNL, the increase number of parameters does not justify accounting for the scale parameters based on BIC and CAIC. However, AIC favors the simple GMNL model, especially compared with SMNL and MNL models.

Moving from MNL to SMNL, only slight improvement in model fit could be observed at the cost of one extra parameter for SMNL, and the two models perform almost identically, with an improvement in one point of the LL at the cost of the SMNL using one extra parameter. That is despite the significant of the parameter of τ . In other words, the results highlighted that for the SMNL, even by adding extra parameters, the fit of the model could not surpass the standard MNL model, highlighting there is no need to account for the scale heterogeneity in the SMNL compared with the MNL model.

The correlations across random parameters were considered for two models, the MIXL and the GMNL models. To conserve space the correlation across the random parameters for both MIXL and GMNL were excluded. However, a few points are worth mentioning about the two implemented techniques. First, pairwise correlation was significant for the GMNL model's random parameters except for correlations across six random parameters such as driver belt status, time of 7:30–10:30 a.m., and van and pickup truck.

Table 2. Model parameters estimates of considered models.

Attr.	MNL		MIXL				GMNL				S-MNL	
			Uncorrelated Error		Correlated Error ""		Uncorrelated Error		Correlated Error ""			
	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.	Est.	St. Err.
Sunny	−0.21	0.060 **	−10.39	6.309	−2.40	1.263	−20.47	6.874 **	−1.00	1.287	−0.39	0.091 **
sd. sunny			10.67	5.510 *	20.11	5.796 **	2.22	1.122 **	−10.01	3.095 **		
11:30–1:30	−0.25	0.070 **	−11.52	6.512 ′	−8.73	2.688 *	−17.82	7.636 *	−6.99	2.362 **	−0.36	0.092 **
sd. 11:30–1:30			1.57	1.776	−1.86	1.505	9.58	4.189 *	3.83	1.621 *		
SUV	0.06	0.115 **	17.14	9.805 ′	7.93	3.309 *	23.72	9.361 *	6.73	3.403 *	−0.14	0.082
sd. SUV			49.61	26.109 ′	9.39	3.478 **	13.14	5.474 *	−21.75	6.069 **		
Pickup truck	0.61	0.066 **	63.91	33.642 *	33.27	9.310 **	47.60	19.108 *	31.01	8.198 **	0.09	0.135
sd. pickup truck			54.48	28.555 *	−0.76	1.738	27.96	11.865 *	−19.29	5.250 **		
7:30–9:30	0.09	0.078	10.12	5.908 ′	5.82	2.689 *	9.46	4.392 *	3.32	1.834	0.73	0.083 **
sd. 7:30–9:30			5.32	7.351	−16.33	4.836 **	3.15	1.803	−16.02	4.427 **		
Driver gender	0.60	0.052 **	16.44	9.131 ′	−13.14	3.736 **	12.94	5.941 *	−7.42	3.595 *	0.05	0.101
sd. Driver gender			110.30	57.580 *	−42.80	11.804 **	77.99	30.769 *	47.19	12.918 **		
Driver’s belting status	0.24	0.076 **	22.73	11.606 *	12.25	3.677 **	16.95	6.068 **	7.73	2.592 *	0.74	0.072 **
sd. Driver’s belting status			41.85	23.929 ′	−27.17	7.834 **	31.21	11.681 **	−19.74	5.815 **		
Day of a week	−0.11	0.065 **	6.23	3.509 ′	6.98	2.292 **	10.86	4.305 *	2.71	2.076	0.29	0.093 **
sd. Day of a week			40.55	21.397 ′	4.27	1.622 **	7.33	3.161 *	−10.28	3.118 **		
Vehicle plate registration	−0.86	0.043 **	−66.91	34.580 *	−20.21	5.652 **	−51.84	20.813 *	−19.23	5.227 **	−0.99	0.065 **
sd. Vehicle plate registration			1.35	1.009	4.33	1.355 **	1.46	0.722 *	1.22	0.407 *	−1	0.087 **
Van	−0.41	0.072 **	−29.56	15.685 ′	−17.08	4.851 **	−17.77	7.000 *	−11.42	4.515 *	−0.48	0.087
SD. van			27.79	15.334	12.21	3.460 **	15.15	5.350 **	−11.06	3.002 **	−0.48	0.087 **
τ			————		————		−0.01	0.003 **	0.07	0.029 *	0.05	0.091
γ^*			————		————		23.18	6.400 **	−3.81	1.560 *	————	————
δ_{sunny}			————		————		−0.01	0.003 **	0.04	0.022 *	−0.26	0.079 **
LL	−3.810		−3701		−3612		−3697		−3618		−3809	
BIC	7725		7577		7795		7596		7833		7725	
AIC	7644		7442		7354		7440		7371		7643	
CAIC	7686		7509		5575		7518		7526		7687	

' statistically significant at 0.1 level; * statistically significant at 0.05 level; ** statistically significant at 0.01 level. "" To conserve space the correlation across the random parameters were excluded.

On the other hand, for the MIXL, the number of significant correlations across pairwise random parameters reduced to 5. Although the AIC favors the correlated GMNL, and MIXL over parsimonious types of these models, measures of BIC and CAIC favor the standard of those models. In general, the more parsimonious models, which did not account for correlations across random parameters, outperform the correlated models.

In summary, all fit measures favor the standard mixed model except for AIC measures. Also, based on the BIC, AIC, and CAIC, the results highlighted that adding scale heterogeneity by SMNL is not needed compared to the standard MNL model. Thus, it can be concluded that BIC and CAIC, similar to the previous study, are more reliable measures for checking for the presence of the scale heterogeneity.

The results highlighted that although accounting for scale heterogeneity is important due to the significance of the tau, its contribution in model fit improvement is not as much to cover up the increased number of parameters. Again, that is despite the fact that magnitude of the tau acknowledges the significance of scale heterogeneity.

The literature review highlighted that accounting for scale heterogeneity through either GMNL or SMNL would likely result in model fit improvement compared with the MIXL [5]. The necessity of accounting for scale heterogeneity might be linked to extreme differences or variation across the observations that the standard mixed model could not account for. However, for our case study we found that adding the scale heterogeneity for giving more flexibility to the distribution is not required. Again, that is despite the significance of the tau parameter and slight improvement in the model fit.

As discussed in the literature review, the justification of a better performance of the GMNL over the MIXL could be due to the structure of the heterogeneity being more complex, which might be linked to a large number of parameters [12]. For our case study, it could be said that our individual-specific data of seat belt use might be not complex enough to necessitate the requirement of the scale heterogeneity in the model.

As the simple mixed model outperforms all other models, we will go over the results very briefly. The standard deviation of all the parameters were included and found to be significant. Those include variables such as sunny weather condition, SUV and pickup type of vehicles, vehicle registration and time of a day. It was interesting to see, although we only considered front-seat passengers seat belt status as the response, that driver gender and driver's belting status were found to be important for the choice of the front-seat passengers seat belt use.

5. Conclusions

Accounting for the individuals' heterogeneity is important to identify reliable and unbiased point estimates of the parameters. That is especially important for evaluation of the choice of vehicle occupants in the use of seat belt as the choice would impact the road fatality dramatically. Various approaches have been taken in the literature review to account for various aspects of heterogeneity to come up with most accurate estimates of the parameters. Those include accounting for scale or heterogeneity in the individual coefficients through taste heterogeneity, or a combination of both. Also considering the correlation across random parameters has often been found to add to the model goodness of fit.

Although it has been argued that the impacts of scale and taste heterogeneities could not be separated, and accounting for scale heterogeneity only gives more flexibility to the distributions, we evaluated six models to provide a vision about the necessity of accounting for various aspects of the individuals observations including heterogeneity in taste, scale or correlations across random parameters error terms. It should be noted that although it is expected that the GMNL would outperform the other techniques, there is no consensus on a preferred technique.

We considered the change in scale heterogeneity based on various attributes such as an intercept and various explanatory variables. The results highlighted that a better fit

would result from considering the scale parameters to vary based on weather condition of sunny. Also, it should be noted that considering several of the explanatory variable for changes of the scale parameter resulted in a lack of convergence. Thus, we concluded that to explain why the scale differs across front-seat passengers, letting the scale to be a function of weather condition would result in the best fit model.

Our results highlighted that although the GMNL provides an improvement based on the log likelihood, the improvement is not as much to cover the added number of parameters due to accounting for the scale heterogeneity. In other words, while AIC, with lower penalty for added parameters, highlighted a small improvement for the GMNL over the MIXL model, BIC and CAIC favored the MIXL model. That is despite the fact that the results highlighted that indeed significant scale heterogeneity exists, but the amount is very small.

In terms of correlated and uncorrelated models, AIC favored a full version of the model with correlated error terms as it imposed a smaller penalty for the number of models' parameters. However, BIC and CAIC still preferred more parsimonious models with no correlation across the random parameters.

A comparison across the standard MNL and SMNL provide a very helpful insight by highlighting that the two models perform almost identically. This indicates that the scale heterogeneity is only trivial, and it is not worth accounting for at the expense of extra numbers of parameters: no further flexibility to the distribution is needed. In summary, considering all incorporated models, AIC favors the parsimonious GMNL model with no correlation, while BIC and CAIC favor parsimonious versions of the MIXL model.

As noted by [5], the question of why sometimes accounting for the scale heterogeneity in the GMNL model performs better than the MIXL model is related to what behavioral pattern the GMNL could explain that the MIXL could not. The scale heterogeneity is needed when the choices of the individuals' respondents are extreme, and the analysis needs an extra layer to account for that extreme behavior. However, that scenario would not apply to our case study, and the mixed model is suitable enough to account for the heterogeneity. That was also acknowledged by a comparison across the SMNL and simple MNL models: the two models perform almost identically in terms of goodness of fit. The results of this study are specific to the dataset being used, and the majority of previous studies acknowledged the improvement in model fits by considering the scale heterogeneity through giving more flexibility to the distribution.

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

Funding: This research funded by 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 copyrighted in 2018. Matching fund for this study was provided by the Mountain-Plains Consortium (MPC).

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

References

1. Seat Belts. Available online: <https://www.nhtsa.gov/risky-driving/seat-belts> (accessed on 1 February 2021).
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] [PubMed]
3. 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]

4. McFadden, D.; Train, K. Mixed MNL models for discrete response. *J. Appl. Econom.* **2000**, *15*, 447–470. [[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. Hess, S.; Rose, J.M. Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation* **2012**, *39*, 1225–1239. [[CrossRef](#)]
7. Scarpa, R.; Thiene, M.; Train, K. Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps. *Am. J. Agric. Econ.* **2008**, *90*, 994–1010. [[CrossRef](#)]
8. Hossain, I.; Saqib, N.U.; Haq, M.M. Scale heterogeneity in discrete choice experiment: An application of generalized mixed logit model in air travel choice. *Econ. Lett.* **2018**, *172*, 85–88. [[CrossRef](#)]
9. Rambonilaza, T.; Brahic, E. Non-market values of forest biodiversity and the impact of informing the general public: Insights from generalized multinomial logit estimations. *Environ. Sci. Policy* **2016**, *64*, 93–100. [[CrossRef](#)]
10. 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 1 February 2021).
11. Louviere, J.J.; Eagle, T. Confound it that pesky little scale constant messes up our convenient assumptions. In Proceedings of the Sawtooth Software Conference, Sequim, WA, USA, January 2006; pp. 211–228.
12. Keane, M.; Wasi, N. Comparing alternative models of heterogeneity in consumer choice behavior. *J. Appl. Econ.* **2012**, *28*, 1018–1045. [[CrossRef](#)]