



OPEN ACCESS

EDITED BY

Francisco Alonso,
University of Valencia, Spain

REVIEWED BY

Mireia Faus,
University of Valencia, Spain
Luis Valentín Montoro,
University of Valencia, Spain

*CORRESPONDENCE

Mahdi Rezapour,
rezapour2088@yahoo.com

SPECIALTY SECTION

This article was submitted to
Transportation and Transit Systems,
a section of the journal
Frontiers in Built Environment

RECEIVED 18 June 2022

ACCEPTED 11 July 2022

PUBLISHED 08 August 2022

CITATION

Rezapour M and Ksaibati K (2022),
Hybrid random utility-random regret
model in the presence of preference
heterogeneity, modeling
drivers' actions.
Front. Built Environ. 8:972253.
doi: 10.3389/fbuil.2022.972253

COPYRIGHT

© 2022 Rezapour and Ksaibati. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Hybrid random utility-random regret model in the presence of preference heterogeneity, modeling drivers' actions

Mahdi Rezapour^{1*} and Khaled Ksaibati²

¹Independent researcher, Marlborough, MA, United States, ²Wyoming Technology Transfer Center, Laramie, WY, United States

Despite the importance of drivers' actions and behaviors, the underlying factors to those actions have not received adequate attention. Understanding the factors contributing to various drivers' actions before crashes could help policy makers to take appropriate actions to tackle those behaviors before crashes occur. One of the first steps could be to identify contributing factors to drivers' actions by using a reliable statistical technique. It is reasonable to assume that drivers vary in their decision-making processes. Thus, in this study, in addition to the random utility maximization (RUM), the random regret minimization (RRM), as a psychological representation of the choice-making process, was considered. While most of the past studies, in the context of traffic safety, focused on either the RRM or RUM, both models' frameworks as hybrid models might be needed to account for the heterogeneity of drivers' decision-making behaviors. In addition, we accounted for the additional dimensions of preference heterogeneity in the latent class (LC) that the model might not capture. The results showed a significant improvement in the model fit of the mixed hybrid LC model compared with the standard hybrid and simple mixed RRM and RUM models. The emotional conditions of drivers, distraction, environmental conditions, and gender are some of the factors found to impact drivers' choices. The results suggest that while the majority of attributes are processed according to the RUM, a significant portion of attributes are processed by the RRM. The hybrid model provides a richer understanding regarding factors to drivers' actions before crashes based on different paradigms.

KEYWORDS

random regret minimization, random utility maximization, hybrid model, drivers' actions, crash, transport

Introduction

Driving cars require constant decision-making processes based on drivers' evaluation of various environmental and roadway characteristics, which could, in turn, be a function of factors such as various demographic and psychological conditions of drivers. Understanding the underlying factors that impact drivers' behaviors is essential to

enhance roadway safety. This is especially important due to the high costs of crashes, where annually, more than 1.2 million people lose their lives as a result of road crashes worldwide (World Health Organization, 2015).

Drivers' behaviors have been defined as the way that the road users choose to or are habitually inclined to drive a vehicle (Öz, Özkan, & Lajunen, 2014). Those behaviors and actions account for more than 94% of the causes of crashes (Singh, 2015). As a result, efforts have been made to understand drivers' actions and behaviors before the involvement in crashes, so appropriate countermeasures could be made to improve the roadway safety. The first step in unlocking factors to drivers' actions and behaviors is by conducting valid statistical techniques.

While doing the statistical analysis, due to the complexity of drivers' choices, it is important not only to account for the data complexity by considering a right statistical technique but also using a right decision-making approach to capture the underlying choice-making processes. It should be noted that here the term "choice" was used as drivers made actions as means of crash avoidances. This is especially important as drivers have a large number of internal cognitive states, where each has its own associated control and transition probabilities (Pentland and Liu, 1999). Also, there is reasonable evidence that choices being made by an individual might be impacted by anticipated regret, and those anticipations would be influenced by the possible outcome of unchosen alternatives (Zeelenberg, 1999). In the regret theory, the individuals engage in a comparative evaluation of the outcome, where instead of maximization of a single alternative, regret is experienced when a choice outcome is worse than what they would have received by choosing a different choice (Connolly and Zeelenberg, 2002).

The selection of various alternatives could influence individuals' decision-making behaviors when they might feel the regret of choosing the wrong decision (Simonson, 1992), or when a nonchosen alternative results in a more desirable outcome than a chosen one (Loomes & Sugden, 1982). Also, the associated regret of selection of various alternatives motivates some individuals to choose some alternatives in favor of others. On the other hand, it has been highlighted that there is a complex effect of emotion on decision-making and reasoning influencing that regret (Blanchette & Richards, 2010).

However, the selection of various choice paradigms depends on the personalities and perception of individual drivers regarding the consequences of taking various actions, and regret alone might not be able to account for all drivers' choice-making behaviors. Thus, it is expected that drivers perform differently in various choice-making behaviors, and while some drivers weigh the consequences of taking alternatives or of potential regret, others care more about the maximization of the utility based on alternatives attributes. Thus, it is important to accommodate those two paradigms in a single-model structure to account for both behaviors.

Thus, in this study, the hybrid technique is adopted to capture heterogeneity across drivers' decision-making process.

The hybrid model could capture both decision-making processes in a single-model framework, while accounting for the preference heterogeneity by the membership in a distinct class of the latent class (LC) model. In this study, it is hypothesized that the individual drivers take various processing steps to take actions resulting in crashes. For instance, those action could follow the random regret minimization (RRM), by pairwise comparison across various alternatives, or solely maximization the utility, random utility maximization (RUM).

Despite the importance of accounting for the presence of various decision-making processes before crashes, no study conducted in the context of traffic safety studies has accounted for both decision-making processes in a single-model framework.

Thus, the main motivation of this research is to identify not only factors influencing various drivers' actions but also to account for various choice-making processes that might be made before crashes. This is especially important as drivers' behaviors account for the majority of crashes. This pattern is expected due to the variation across drivers. Some attributes before crashes might be processed by drivers through the RUM, while others with the RRM. As such, considering the hybrid model could provide new insights into the ways that drivers process attributes before crashes. This is especially important as it has been found that there is a confounding between heterogeneities in sensitivities and in decision rules (Hess, Stathopoulos, & Daly, 2012).

In summary, the main contributions of this study in terms of parameters estimations are: 1) in addition to accounting for the preference heterogeneity, captured by membership of the distinct class of the LC model, we accounted for the heterogeneity in the decision rule by using a hybrid model and 2) we included another layer of preference heterogeneity within a class of the LC model to capture the heterogeneity that the standard LC model might not be able to capture. All these considerations are expected to result in more unbiased estimates of parameters.

Literature review

This literature review presents a number of studies conducted studying drivers' actions. It will then mainly discuss those studies that implement the RRM as a single-model framework, and its extension to the hybrid model.

Drivers' actions and behaviors

A study investigated the drivers who choose to speed and estimated the time saved by drivers who took that action (Peer, 2011). The bias in the estimated time was linked to the drivers' estimation of the required speed. In another study, driver behavioral parameters impacting the choice, by means of a

questionnaire-based survey, was conducted (Yannis, Kanellou, Aggeloussi, & Tsamboulas, 2005). The results of the logistic model highlighted the parameters, related to trip duration and increased trip, impacting the choice of risk reduction alternative. Also, it was found that driver characteristics such as gender and income impact the crash risk reduction choice.

It has been observed that various drivers' actions, such as lane-keeping behavior, are a function of parameters that drivers observed (Liu & Salvucci, 2001), and also personal characteristics of drivers (Toth-Szabo, Kovacs, & Horvath, 2006). A survey-based technique was used for modeling and recognizing the drivers' behaviors (Wang, Xi, & Chen, 2014). A wide range of drivers' characteristics were used to improve vehicle's dynamic performance and decrease driver's workload.

In another study, factors to youthful driving behavior were investigated and grouped into six factors (Shope, 2006). The results highlighted that driving ability, development, personality, demographics, perceived environment, and driving environment are some of the factors influencing the drivers' actions. In another study, factors affecting driver's speed selection was investigated (Liu et al., 2020). The results highlighted that driver speed behaviors would be impacted by factors such as gender, driving experience, and number of night-driving per weeks.

In another study, only drivers actions in case where vehicles has some forms of malfunctions at the time of crash was considered (Rezapour & Ksaibati, 2022). In the model, due to the endogeneity bias, vehicle malfunctions were indirectly modeled by means of latent variable (LV). The results, for instance, highlighted that female driver are less likely to be involved in speeding driving due to various driving actions, such as brake defect. Also, driving under the influence of drug and alcohol were some of important factors, impacting various drivers' actions.

Random regret minimization and hybrid models

This section will outline studies applying the RUM and RRM models and it ends with studies employed the hybrid technique. The RRM, as a plausible behavioral paradigm for studying the crash avoidance maneuvers, was employed (Kaplan & Prato, 2012). Only single-layer models of the RUM and RRM were employed. A comparison across the standard RRM and RUM highlighted a better fit of the RRM model compared with the RUM counterpart.

On the other hand, extensive efforts have been made implementing either regret or utility-based decision in other areas of transportations [e.g., (Hensher, Greene, & Chorus, 2013)]. However, it has been highlighted that there is a confounding between taste heterogeneity and decision rule heterogeneity (Hess et al., 2012), so it is important to

accommodate different decision rules in a hybrid structured model to reduce the confounding between those two factors. Studies have been conducted with the help of separate structured models and hybrid model to analyze various aspects of choice-making processes in the literature. Two separate models of the RRM and RUM, while accounting for the preference heterogeneity, were implemented (Hensher et al., 2016). The results highlighted a superiority of the mixed single-layer models compared with the standard models. Also, the RRM and RUM perform almost equally in terms of the goodness of fit.

The importance of accounting for differences in the choice process of individuals has been discussed in the literature (Hess et al., 2012). In the study, the model's implementation was based on the LC hybrid model, combining both the RUM and RRM. A better fit of hybrid model compared with a single-model framework of the RRM and RUM was observed.

A hybrid of RRM and RUM, while accommodating the taste heterogeneity, was used to evaluate the stated choice in the context of calming schemes (Thiene et al., 2012). The hybrid mixed model outperformed the other considered model. However, it should be noted that due to a high complexity of the model, only a limited number of parameters were included in that study.

In another study, satisfaction and uncertainty in car-sharing decisions with the help of the random regret-based hybrid choice model was investigated (Kim et al., 2017). The study did not compare any other plausible model framework, and only the standard hybrid model was used.

The hybrid model, in another study, was employed on two datasets (Chorus et al., 2013). The results highlighted that in terms of willingness-to-pay (WtP) measures, the hybrid model differs substantially from the conventional utility-based model. Thus, allowing for heterogeneity across the processing rules would lead to a significant improvement in the model fit.

The mixed random utility-random regret model was used, where class allocation was driven by a latent variable (Hess & Stathopoulos, 2013). The results highlighted that the most regret-prone respondents align their real-life commute with their aspirational values. Also, it was discussed the importance of implementation of the hybrid model due to confounding between the decision rule heterogeneity and simple heterogeneity in sensitivities.

Although studies have contributed to the body of knowledge by investigating various factors influencing drivers' actions, those studies did not account for the preference heterogeneity, or heterogeneity in choice-making process. Also, limited studies implemented the single structured model of the RUM in the context of traffic safety. However, as discussed, not necessarily all the drivers use the same choice paradigm while making action, while some use the RUM, others might use the RRM, so it is important to account for those variations in a single-model framework. On the other hand, another limitation of the past studies was not accounting for the heterogeneity across drivers in terms of making various

actions before crashes, which might result in biased or even erroneous point estimates. Thus, this study was conducted to account for those limitations by considering both the RUM and RRM paradigms in a hybrid LC model.

Materials and methods

As the hybrid LC model combines the two decision-making frameworks of the RUM and RRM, we will first briefly outline the two techniques and then we will discuss the aggregation of the two techniques in a single LC model.

The total utility of an alternative i and individual driver n could be written as

$$U_{ni} = f(\beta, x_{ni}) + \varepsilon_{ni}, \tag{1}$$

where x_{ni} is a vector of an attribute related to an alternative i and individual n , where same attributes β were used across all alternatives. β is a utility coefficient to be estimated and ε_{ni} is the type I error.

Both the RRM and RUM are based on the multinomial logit model (MNL), highlighting the probability of individual n for choosing alternative i over all alternatives of j as (McFadden, 1973).

$$Pr(RU_{ni}) = \frac{e^{\beta x_{ni}}}{\sum_{j=1}^J e^{\beta x_{nj}}}. \tag{2}$$

Now in the context of the RRM, while making a choice, it is postulated that some drivers would make a pairwise comparison across all possible alternatives with an objective of minimizing the regret before making a decision. That is,

$$\hat{R}_{ni} = R_{ni} + \varepsilon_{ni} = \sum_{j \neq i} \sum_m \ln(1 + \exp(\beta_m^j \cdot x_{njm} - \beta_m^i \cdot x_{nim})) + \varepsilon_{ni}, \tag{3}$$

where R_{ni} is the deterministic part and ε_{ni} is the stochastic component. m and n are the indices of attribute and individual, respectively, and i and j are two arbitrary alternatives. In the abovementioned equation, as we have only alternative-specific attributes, similar vectors of x_{njm} and x_{nim} , for two arbitrary alternatives would be incorporated. Eq. 3 could be compared with $\max(0, \beta_m^j \cdot x_{njm} - \beta_m^i \cdot x_{nim})$, which postulates that the regret equals zero when alternative i perform better than another alternative j for an attribute of m . However, the formula would be problematic when $\beta_m^j \cdot x_{njm} - \beta_m^i \cdot x_{nim}$ equates to zero, resulting in nonsmoothed likelihood function, triggering the application of Eq. 3 (Chorus et al., 2013). Also based on Eq. 3, the regret would be minimized as the variation between $\beta_m^j \cdot x_{njm} - \beta_m^i \cdot x_{nim}$ is minimal and would increase when the variation between the two increases. In Eq. 3, in case of the zero differences between the two attributes, the related \hat{R}_{ni} would be around 1, which is considered as rejoice.

Now, the objective would be the minimization of the regret or the maximization of the negative of regret as

$$Pr(RR_{ni}) = \frac{\exp(-\hat{R}_{ni})}{\sum_{j=1}^J \exp(-\hat{R}_{nj})}. \tag{4}$$

Similar to the mixed RUM, the RRM could be expanded to the mixed model like

$$P_{ni} = \int \frac{\exp(R_{ni})}{\sum_j \exp(R_{nj})} f(\beta|\varphi) d\beta. \tag{5}$$

Now, the final model, based on the LC model, could be written based on the probability of the choice and membership probability to shape the mixed LC model, where membership probability would be written as

$$\pi_s = \frac{\exp(\alpha_s + \gamma_s z)}{\sum_{s=1}^S \exp(\alpha_s + \gamma_s z)} \text{ Where } \sum_{s=1}^S \pi_s = 1 \text{ and } 0 \leq \pi_s \leq 1, \tag{6}$$

where γ_s and z would be used if we want our class membership to be set based on a set of observed attributes, where γ_s , in addition to α_s , will be estimated. However, in this study, we just set the class membership based on the constant values of α_s . Also, for identification, for an arbitrary class, α_s is constrained at zero. Now, the latent class model combining the class probability with choice probability could be written as

$$Pr(y) = \sum_{s=1}^S \pi_s P_{ni}. \tag{7}$$

Mathematically breaking down the individual's drivers into two behavioral processes of the RRM and RUM, considering the preference heterogeneity with a class, could be written as

$$Pr_n = \pi_{RR_n} Pr(RR_n) + \pi_{RU_n} \int Pr(RU_n) f(\beta|\Omega) d\beta, \tag{8}$$

where the final likelihood is the sum of all probability related to the two behavioral processes across all observations of n and β is the parameter of random parameter following normal distribution with the parameters of Ω . We did not incorporate and include random parameter for the RRM part of model as no attribute was found to be important to be considered as random parameter. Based on Eq. 8, we only accounted for the heterogeneity in the RUM layer for a single parameter due to the high number of incorporated parameters and lack of convergent. Again, we considered the preference heterogeneity for some attributes of the RRM, which either did not result in a significant standard deviation for that parameter distribution or resulted in a lack of convergence. It should be noted that normal distributions were considered for all random parameters.

The model parameters estimations are based on Eq. 8, by computing the within-class choice probabilities by means of MNL and the probabilities of the latent class models, π , and optimizing the two parts simultaneously. The core part of the

TABLE 1 Descriptive summary of attributes.

Drivers' actions	Categories	Aggressive driving actions	Failure to keep proper lane	Improper passing	Others	No improper driving
Driver's distraction	Normal ^a	874 (1.2)	3,107 (4.4)	336 (0.47)	24,738 (35)	23,367 (33)
	Distracted	1,194 (1.7) ^b	1,965 (2.8)	113 (0.16)	13,587 (19)	1,565 (2.2)
Condition of drivers	Normal ^a	1,181 (1.7)	3,018 (4.1)	360 (0.51)	29,326 (41.3)	24,501 (34.5)
	Others	887 (1.2)	2,054 (2.9)	89 (0.5)	8,999 (12.7)	431 (0.61)
Driver license registration	Wyoming resident ^a	1,530 (2.2)	2,837 (4.0)	276 (0.4)	25,473 (23.4)	16,637 (35.9)
	Others	538 (0.76)	2,235 (3.15)	173 (0.25)	12,852 (11.7)	8,295 (18.1)
Gender	Male ^a	1,431 (2.02)	3,553 (5)	309 (0.44)	25,092 (35.3)	(22.5)15,948
	Female	637 (0.90)	1,519 (2.2)	140 (0.20)	13,233 (18.63)	8,984 (12.65)
Roadway condition	Others ^a	1,880 (2.7)	3,935 (5.5)	400 (0.54)	26,752 (38)	21,568 (30.36)
	Icy	188 (0.27)	1,137 (1.6)	49 (0.07)	11,573 (16.3)	3,364 (4.70)
Lighting condition	Light ^a	1,417 (2)	3,131 (4.41)	355 (0.5)	26,204 (36.89)	11,420 (16.08)
	Dark	651 (0.92)	1,941 (2.76)	94 (0.13)	12,121 (17.06)	13,512 (19.02)
Vehicle type	Others ^a	1,966 (2.77)	4,321 (6.08)	405 (0.57)	35,100 (49.42)	23,494 (33.076)
	Trucks	102 (0.14)	751 (1.06)	44 (0.06)	3,225 (4.54)	1,438 (2.02)

^aReference category used in the analysis.

^bFrequency (%).

model is defining the likelihood functions for each alternative related to the RUM and RRM. It is clear that observations are assigned to different paradigms, latent class, as drivers employed different choice-making paradigms, before making various drivers' actions. Also, the observations' allocation to different classes, being related to π_{RR} and π_{RU} part of Eq. 8, are implemented by maximizing the sum of all observations related to Pr_n in Eq. 8, through maximum likelihood process.

Data

The data in this study was obtained directly from the Wyoming Department of Transportation (WYDOT) during a 5-year period, 2015–2019. Primarily, four datasets were aggregated for the final dataset. Those included: crash, vehicle, driver, and environmental characteristics datasets. The data were aggregated based on the common columns of ID, milepost, and highway system. Our data accounted for the total of 39,934 observations (drivers' actions) before crashes.

In this study, we focus only on the drivers' at-fault maneuvers. As the exact maneuvers or actions of the drivers resulted in crashes were unobserved, the actions were determined by the highway patrol at the scene of crashes based on the available circumstances. As could be expected from this type of data, all attributes were individual-specific and no choice-specific attributes were included in the model. The choices included in this study were: 1) aggressive driving, which includes erratic, reckless, or aggressive types of driving; 2) failure to keep proper lane; 3) improper passing; and 4) if drivers were not involved in

any type of improper actions, the action was coded as “no improper driving.” Also, if the driver actions did not come under categories of any of the aforementioned actions, it was categorized under “others”.

Distraction variable in the data is defined as any interruption in the cabin such as electronic communication (for example, cell device), a passenger or pet in the cabin, or even distraction outside the motor vehicle. The driver conditions at the time of crash include being fatigued, fainted or asleep, or emotional conditions such as sadness, depression, or anger, compared with the normal driving condition as a reference category. The descriptive summary of the important predictors is presented in Table 1.

Results

The results section is presented in two subsections. First, the results of goodness of fits of the considered models is presented. Then, the results of the best fit of the mixed LC model, along with the standard LC model for a comparison purpose is presented.

Model comparisons

Although the hybrid mixed model could not exactly be considered as an extension of the mixed RRM or RUM, to gain a better perspective regarding the model performances, the model fits of those models are also presented. A total of four models were considered and compared. As the standard mixed RRM and RUM could take only a single layer and, here,

TABLE 2 Comparison across considered models.

ID	Models	No. of parameters	LL	AIC	BIC	Rho-square
1	Mixed RRM	22	-65,656	131,355	131,558	0.42
2	Mixed RUM	22	-65,504	131,053	131,254	0.43
3	Standard hybrid model	34	-63,508	127,083	127,432	0.44
4	Hybrid mixed model	38	-60,931	121,938	122,218	0.47

there are more parameters in the second layer of the LC model with less constrained parameters, we considered the second layer of those models for model 1 and 2, see [Table 2](#).

It can be seen by considering only the single layers that the RRM model slightly outperforms the RUM model. However, both models fall short when it comes to the standard LC model. Moving to the LC model considering the additional random taste heterogeneity, better performance could be observed. The gap between Akaike information criterion (AIC) and Bayesian information criterion (BIC) for all the considered models are so wide that it does not make a difference which measure is used for models' comparisons.

In summary, the results highlighted that adding the continuous preference heterogeneity on top of the RUM in the LC model outperform all the other considered models, highlighting that there is still heterogeneity in the model's components that solely the LC model might not fully account for. The Rho-square as a measure of an improvement in the log likelihood of the starting log likelihood to the final likelihood is also presented to provide some insight. It should be reiterated that the model fit is not the only criteria driving the modelling process, but the main goal of implementing the hybrid model was providing a better understanding regarding the choice-making process of drivers prior to crashes.

Model results

Due to the lack of observations for some attributes, and lack of convergent, several parameters are fixed at their starting values. Constraining the attributes are especially important, as initially there were a large number of parameters (+60), and a lack of convergent is expected, especially when adding random parameters on top of the LC model. The lack of convergence for the mixed LC model forces the researchers not to include many parameters or skipping the inclusion of random parameters in the LC model [(see for instance ([Boeri et al., 2014](#)))]. Various driver actions as alternatives were considered in this study. These include short-term behaviors/alternatives, such as passing, compared with long-term behaviors, such as lane keeping, or a combination of both behaviors such as aggressive driving.

[Table 3](#) presents the two models that outperformed the other considered models, including the LC and the mixed LC models, as shown in [Table 2](#). It should be noted that, due to having all individual-specific observations and for the identification purposes, "others" type of drivers' action was normalized to include only the constant value, so the results of that alternative are not presented in [Table 3](#). Here, we only considered a constant value for "others" instead of "no improper driving" for several reasons. First, "others" is a combination of various drivers' actions and factors to this drivers' actions might not be of interest to the policy makers compared with other drivers' action as they account for a variety of actions. Second, it is important to highlight factors to "no improper driving actions" as this action accounts for a significant proportion of crashes, and it is important to understand factors contributing to crashes when the drivers did not involve in any improper action.

Based on the discussed formulations in the methodology section, the interpretations of the estimated parameters of RRM and RUM vary. For instance, the negative sign of the coefficient of gender in the regret component in [Table 3](#) indicates that the regret of the choice of aggressive driving and no improper driving decrease when the level of attribute changes from 0 (male) to 1 as female, so this alternative would be less likely compared with all the other alternatives, while drivers making a choice.

As can be seen from [Table 3](#), there are instances where the coefficient estimates are discordant in terms of significance, and the directions of sign across the RRM and RUM. For instance, while driver distraction is not different than zero for aggressive driver action for the RUM model, it is significant in the RRM model for the same alternative. Additionally, while gender has a positive sign for the RRM for the alternative of no improper driver action, the sign is negative for the same alternative across the RUM component. These differences are expected due to the heterogeneity across individuals' drivers in terms of choice-making behaviors. However, it should be noted that most of the attributes have consistent signs across the two components.

The changes could also be observed across the LC and the mixed LC model. For instance, while the estimate of driver condition of aggressive driving is significant for the standard RUM model, it is not different from zero when we accounted for the taste heterogeneity. The same report holds true with the condition of drivers. Also, when the estimated coefficient of

TABLE 3 Latent class RUM–RRM model with and without taste heterogeneity, t-ratio in parentheses.

Attributes	Accounting for taste heterogeneity				No taste heterogeneity			
	Aggressive driving actions	Failure to keep proper lane	Improper passing	No improper driving	Aggressive driving actions	Failure to keep proper lane	Improper passing	No improper driving
RRM								
Driver distraction	1.346 (16.07)	1.168 (14.95)	0.340 (4.43)	—	1.00 (10.835)	1.05 (11.621)	0.059 (0.665)	—
Driver conditions	0.401 (4.258)	—	—	-1.929 (-28.017)	0.563 (5.99)	—	—	-2.162 (-20.5)
Driver license registration	-1.284 (-23.5727)	—	-1.247 (-9.558)	0.490 (16.530)	-1.419 (-25.73)	—	-1.208 (-26.89)	0.295 (7.051)
Gender	-0.165 (-1.926)	—	-0.148 (-1.729)	0.552 (9.192)	-0.143 (-1.520)	—	-0.0885 (-1.520)	0.630 (8.850)
Roadway condition	—	—	—	—	—	—	—	—
Lighting condition	—	—	-0.699 (-9.740)	—	—	—	-1.084 (-15.18)	—
Vehicle type (truck)	—	—	—	-1.041 (-11.13)	—	—	—	-0.909 (-1.041)
RUM								
Driver distraction	0.152 (0.8)	—	—	-13.842 (-0.42263)	0.186 (1.61)	—	—	-3.161 (-22.0)
Driver conditions	-0.422 (-1.1546)	4.705 (3.575)	—	-30.730 (-0.78898)	-0.679 (-26.88)	-0.0026 (1.71)	—	-3.416 (-17.342)
Driver license registration	-1.846 (-14.472)	-5.662 (-0.328)	-13.707 (-0.629)	0.510 (24.032)	-1.460 (-16.297)	-10.071 (-1.893)	-18.01 (-0.102)	1.078 ()
Gender	-0.386 (-2.752)	-0.111 (-0.562)	-0.148 (-1.729)	-0.371 (-7.888)	-0.9812 (-3.570)	-0.003 (-0.02)	-0.089 (-0.959)	-0.094 (-2.731)
Roadway condition	-10.765 (-0.456)	-19.000 (-2.835)	-0.05 (-0.001)	-16.566 (-2.473)	-14.59 (-0.1)	-8.806 (-0.463)	8.974 (0.051)	-2.258 (-32.8)
Lighting condition	-8.542 (-0.557)	—	—	—	-11.297 (-0.261)	—	—	—
Vehicle type (truck)	-1.109 (-1.076)	0.784 (1.587)	—	-0.208 (-1.918)	-0.670 (-1.778)	0.067 (0.153)	—	-0.957 (-13.244)
Random parameters								
Dark, mean	—	-10.123 (-2.226)	—	9.864 (2.5867)	—	—	—	—
Dark, SD	—	-11.842 (-2.772)	—	-7.786 (-2.327)	—	—	—	—
Null LL: -114,022, final LL:-60,932, # par:38, BIC:122,329 adj Rho-square = 0.47					Null LL: -114,022, final LL:-63,507, # par:34, BIC:127,432 adj Rho-square = 0.47			
Component	Shares	estimate	t-ratio		Component	Shares	estimate	t-ratio
Delta_RRM	0.3	—	—		Delta_RRM	0.26	—	—
Delta_RUM	0.7	0.82	33		Delta_RUM	0.74	1.03	35.66

roadway condition is around -2 for no improper driving in the standard model, when we accounted for the heterogeneity, it decreases to -16.6.

It is also interesting that the impact of distraction for improper passing is not different from zero for the standard model. However, the impact becomes significant for the RRM model when we account for the heterogeneity of the RUM model.

All of the aforementioned points highlight the importance of capturing the extra preference heterogeneity within a class of the LC for highlighting unbiased point estimates of parameters.

Based on the regret component of the model, our results highlighted that those drivers, making actions while being under non-normal condition are more likely to make aggressive driving actions and less likely to make “no improper driver actions”. The

results of the RUM model highlighted that those drivers are also more likely to fail to stay in the proper lane. The importance of specific emotions on judgment and driving performance by car simulation studies was highlighted in a previous study (Jeon, Walker, & Yim, 2014). This study found that induced anger shows negative effects in degraded driving performance and decision-making process.

Those drivers having a distraction in motor vehicle are likelier to make drivers' actions of aggressive driving, failure to keep lane, and improper passing. However, this attribute was not significant and not included for the RUM class of the model, being related to failure to keep proper lane and improper passing. The results highlight the fact that the attributes of distraction for those two alternatives are likelier to be processed by the RRM process.

It is interesting to note that while the nonresidence of Wyoming, due to possible care they are taking, travelling on the mountainous areas of Wyoming, they are less likely to partake in aggressive driving and improper passing. However, due to a possible lack of familiarity with the roadway, they are likelier to make "no improper actions". The estimated parameters have also similar directions for the RUM part of the model. It should be noted that the attribute for failure to keep proper lane was not processed by the RRM model.

Regarding lighting condition for the RRM, all parameters are constrained except for improper passing, highlighting that dark condition decreases the probability of improper passing. This could be due to the fact that drivers drive more cautiously in less-than-optimal conditions. Also, this attribute is the only variable being considered as random for the second class, RUM, of the model for no improper driving and failure to keep proper lane.

The considered random parameters distributions both have significant standard deviation (SD), indicating a significant presence of taste heterogeneity across the individual's drivers in the RUM part of the model. For instance, the predictor lighting condition is an indicator variable for dark (1) versus light (0). The estimate for this random parameter is -10.11 and the estimated standard deviation is 11.842 for the driver action of failure to keep proper lane. This means that a significant majority, 80%, of drivers' action of failure to keep proper lane was made in light condition while only 20% of them were made in dark condition.

It is also noteworthy, from the bottom of Table 3, adding taste heterogeneity to the RUM class reduces the share and includes more observation in the RRM class. This might be due to the biased performance of the RUM capturing more observations, while not accounting for the taste heterogeneity. For our analysis, the RUM dominates in terms of holding more individuals, for both the standard and mixed models. The reason has been linked to the fact that the RRM model is especially important for those individuals who found it difficult to make the right decision (Zeelenberg and Pieters, 2007). As the majority of our observed subjects are residents of Wyoming, and the residents are more familiar with the roadway, it could be speculated that they would

be more likely to make decisions based only on the maximization of their utility. T-ratio is provided as a means of parameters significance which is coefficients divided by the standard error.

Discussion

The choice-making behaviors of drivers could be described as a device with a number of mental states, each with a particular control behavior. These behaviors have been described as a process to describe actions made by drivers for short-term or long-term behaviors (Pentland and Liu, 1999). It is expected that the final choice would be made by a driver based on the part of human control that was active at the time of crashes. Thus, we expect that the behavioral process of individuals' drivers, while making various driving actions, does not necessarily follow the RUM.

As a result, the LC framework was expanded to two paradigms of the RRM and RUM based on their behavioral processes, as a single decision-making process might not be able to capture all the decision-making process. The regret is expected to be experienced by drivers when a nonchosen alternative outperforms the chosen ones, and thus some drivers aim to minimize the anticipated regret.

The current study is one of the earliest studies that investigated the drivers' behaviors in the context of the hybrid model. Also, while considering the LC model, the homogeneity was mainly considered across the attributes within each class. It is worth accounting for the possible heterogeneity that might still exist within each class. Thus, we added another layer of random parameter on top of the LC model to capture the extra heterogeneity that the LC model might not capture. Here, we will discuss some of our findings.

It was found that driver distraction increases the likelihood of all risky driving actions except for no improper driving. The impact of drivers' distraction could be linked to the deterioration of driving skills and performance which result in various driving actions such as aggressive driving. The link between aggressive driving and inattention, for instance, was acknowledged in the previous study (Alkinani et al., 2020).

It was also found that drivers with no normal conditions such as anger or emotional, being processed with the RRM, are likelier to involve in aggressive driving action. The result is in line with the past studies that anger is usually associated with aggressive driving (Nesbit et al., 2007). However, it should be noted that, here, non-normal conditions include other conditions such as emotional conditions of drivers was considered.

The gender variable was processed by both layers and it was found that those drivers have a decreased odd of aggressive driving and improper passing. The result is in line with the past study that, for instance, men are more aggressive than women in terms of crash involvement (Shinar, 1998).

It was found that drivers on less-than-optimal road conditions are less likely to be involved with the failure to keep proper lane. Although the results seem to be counterintuitive, it has been discussed that the reason could be linked to the fact that drivers drive more cautiously on less-than-optimal road conditions (Moomen et al., 2019).

It is important to discuss the study limitations. It could be argued that the current study only employed the individual-specific attributes, and thus the scope of the study might be limited. In addition, the choice is not presented directly to the drivers, but information collected by the highway patrol at the time of crashes. We considered those possible limitations as uniqueness of this study and the fact that all the attributes and choices are obtained from real-life scenarios. Also, in case of the presentation of the scenarios to individuals, there might be biased responses to our question.

A second possible argument is the presented information such as driver license registration are just demographic characteristics of individuals, and thus could not be considered as attributes but features of other unseen characteristics of drivers. Although the point is accepted to some extent, it should be noted that we incorporated all the characteristics even if they highlight other characteristics of drivers. For instance, driver's license attribute might give information regarding drivers' characteristics such as personalities and familiarity with roadways.

An important limitation of this study, being acknowledged by previous studies, was a lack of convergence of the mixed LC model due to the incorporation of a continuous layer of taste heterogeneity on top of the LC model. As a result, the majority of past studies incorporated very limited attribute while incorporating the random heterogeneity in the LC model. Here, we faced a similar issue, and we conducted many trials and error by constraining some attributes to converge our model. For instance, important factors such as speeding, driver experience, and alcohol consumption were dropped from the model, either due to the lack of significance in the prediction of the driver actions or lack of convergence.

Concluding remarks and recommendations

The implemented method provides a richer understanding regarding the choice-making behaviors of drivers before crashes, compared with a simple RUM or RRM. For instance, it was discussed that more attributes would be processed by the RUM rather than the RRM, while drivers making various drivers' actions before crashes. A better fit of the hybrid model might be an indication of necessity of accounting for processing-rule

heterogeneity in addition to taste heterogeneity. Moreover, it has been discussed that the implementation of both paradigms and accounting for the preference heterogeneity within a class of the LC model could offer unbiased point estimates of factors to drivers' behaviors before crashes. That was seen in terms of changes in point estimates and their significances.

By better understanding the factors contributing to various drivers' actions and behaviors, policies and appropriate interventions would be used to amend those actions. The final outcome of interest is changing drivers' actions and consequently the reduction of crashes. Several factors are found to be especially important in choice-making behaviors of drivers. For instance, driver conditions and distraction could be amended by intervention. On the other hand, perceived environment seems to incorporate the most complex set of factors shaping drivers' actions and behaviors. Educating the drivers regarding the mountainous topography of the state might provide a better understanding for drivers regarding the hazard ahead. The law enforcement is also expected to influence drivers' actions and behaviors, which might go unchecked. This could be achieved by restricting the drivers who are not under normal conditions.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

MR: writing. KK: supervision.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Alkinani, M. H., Khan, W. Z., and Arshad, Q. (2020). Detecting human driver inattentive and aggressive driving behavior using deep learning: Recent advances, requirements and open challenges. *Ieee Access* 8, 105008–105030. doi:10.1109/access.2020.2999829
- Blanchette, I., and Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition Emot.* 24 (4), 561–595. doi:10.1080/02699930903132496
- Boeri, M., Scarpa, R., and Chorus, C. G. (2014). Stated choices and benefit estimates in the context of traffic calming schemes: Utility maximization, regret minimization, or both? *Transp. Res. Part A Policy Pract.* 61, 121–135. doi:10.1016/j.tra.2014.01.003
- Chorus, C. G., Rose, J. M., and Hensher, D. A. (2013). Regret minimization or utility maximization: It depends on the attribute. *Environ. Plann. B. Plann. Des.* 40 (1), 154–169. doi:10.1068/b38092
- Connolly, T., and Zeelenberg, M. (2002). Regret in decision making. *Curr. Dir. Psychol. Sci.* 11 (6), 212–216. doi:10.1111/1467-8721.00203
- Hensher, D. A., Greene, W. H., and Chorus, C. G. (2013). Random regret minimization or random utility maximization: An exploratory analysis in the context of automobile fuel choice. *J. Adv. Transp.* 47 (7), 667–678. doi:10.1002/atr.188
- Hensher, D. A., Greene, W. H., and Ho, C. Q. (2016). Random regret minimization and random utility maximization in the presence of preference heterogeneity: An empirical contrast. *J. Transp. Eng.* 142 (4), 04016009. doi:10.1061/(asce)te.1943-5436.0000827
- Hess, S., and Stathopoulos, A. (2013). A mixed random utility—Random regret model linking the choice of decision rule to latent character traits. *J. Choice Model.* 9, 27–38. doi:10.1016/j.jocm.2013.12.005
- Hess, S., Stathopoulos, A., and Daly, A. (2012). Allowing for heterogeneous decision rules in discrete choice models: An approach and four case studies. *Transportation* 39 (3), 565–591. doi:10.1007/s11116-011-9365-6
- Jeon, M., Walker, B. N., and Yim, J.-B. (2014). Effects of specific emotions on subjective judgment, driving performance, and perceived workload. *Transp. Res. Part F Traffic Psychol. Behav.* 24, 197–209. doi:10.1016/j.trf.2014.04.003
- Kaplan, S., and Prato, C. G. (2012). The application of the random regret minimization model to drivers' choice of crash avoidance maneuvers. *Transp. Res. Part F Traffic Psychol. Behav.* 15 (6), 699–709. doi:10.1016/j.trf.2012.06.005
- Kim, J., Rasouli, S., and Timmermans, H. (2017). Satisfaction and uncertainty in car-sharing decisions: An integration of hybrid choice and random regret-based models. *Transp. Res. Part A Policy Pract.* 95, 13–33. doi:10.1016/j.tra.2016.11.005
- Liu, A., and Salvucci, D. (2001). "Modeling and prediction of human driver behavior," in Intl. Conference on HCI, 1479–1483.
- Liu, J., Cai, J., Lin, S., and Zhao, J. (2020). Analysis of factors affecting a driver's driving speed selection in low illumination. *J. Adv. Transp.*, 1–8. doi:10.1155/2020/2817801
- Loomes, G., and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Econ. J.* 92 (368), 805. doi:10.2307/2232669
- McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior*. Berkeley, California: University of California at Berkeley.
- Moomen, M., Rezapour, M., and Ksaibati, K. (2019). An investigation of influential factors of downgrade truck crashes: A logistic regression approach. *J. traffic Transp. Eng. Engl. Ed.* 6 (2), 185–195. doi:10.1016/j.jtte.2018.03.005
- Nesbit, S. M., Conger, J. C., and Conger, A. J. (2007). A quantitative review of the relationship between anger and aggressive driving. *Aggress. Violent Behav.* 12 (2), 156–176. doi:10.1016/j.avb.2006.09.003
- Öz, B., Özkan, T., and Lajunen, T. (2014). Trip-focused organizational safety climate: Investigating the relationships with errors, violations and positive driver behaviours in professional driving. *Transp. Res. Part F Traffic Psychol. Behav.* 26, 361–369. doi:10.1016/j.trf.2014.03.004
- Peer, E. (2011). The time-saving bias, speed choices and driving behavior. *Transp. Res. Part F Traffic Psychol. Behav.* 14 (6), 543–554. doi:10.1016/j.trf.2011.06.004
- Pentland, A., and Liu, A. (1999). Modeling and prediction of human behavior. *Neural Comput.* 11 (1), 229–242. doi:10.1162/089976699300016890
- Rezapour, M., and Ksaibati, K. (2022). Modeling factors to various drivers' actions due to vehicle malfunctions, the integrated latent class and latent variable model. *Eng. Rep.*, e12537. doi:10.1002/eng.2.12537
- Shinar, D. (1998). Aggressive driving: The contribution of the drivers and the situation. *Transp. Res. Part F Traffic Psychol. Behav.* 1 (2), 137–160. doi:10.1016/s1369-8478(99)00002-9
- Shope, J. T. (2006). Influences on youthful driving behavior and their potential for guiding interventions to reduce crashes. *Inj. Prev.* 12 (1), i9–i14. doi:10.1136/ip.2006.011874
- Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *J. consumer Res.* 19 (1), 105. doi:10.1086/209290
- Singh, S. (2015). *Critical reasons for crashes investigated in the national motor vehicle crash causation survey (No. DOT HS 812 115)*.
- Thiene, M., Boeri, M., and Chorus, C. G. (2012). Random regret minimization: Exploration of a new choice model for environmental and resource economics. *Environ. Resour. Econ. (Dordr).* 51 (3), 413–429. doi:10.1007/s10640-011-9505-7
- Toth-Szabo, Z., Kovacs, K., and Horvath, R. (2006). "Human factor in traffic models," in IEEE International Conference on Computational Cybernetics, Tallinn, Estonia, 20–22 Aug. 2006. Paper presented at the 2006.
- Wang, W., Xi, J., and Chen, H. (2014). Modeling and recognizing driver behavior based on driving data: A survey. *Math. Model. Analysis, Adv. Control Complex Dyn. Syst.* 2014, 245641. doi:10.1155/2014/245641
- World Health Organization (2015). *Global status report on road safety 2015*. Geneva, Switzerland: World Health Organization.
- Yannis, G., Kanellopoulou, A., Aggeloussi, K., and Tsamboulas, D. (2005). Modelling driver choices towards accident risk reduction. *Saf. Sci.* 43 (3), 173–186. doi:10.1016/j.ssci.2005.02.004
- Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *J. Behav. Decis. Mak.* 12 (2), 93–106. doi:10.1002/(sici)1099-0771(199906)12:2<93::aid-bdm311>3.0.co;2-s
- Zeelenberg, M., and Pieters, R. (2007). A theory of regret regulation 1.0. *J. Consumer Psychol.* 17 (1), 3–18. doi:10.1207/s15327663jcp1701_3