Analytic Methods in Accident Research 2 (2014) 30–38

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

journal homepage: www.elsevier.com/locate/amar

A latent class analysis of single-vehicle motorcycle crash severity outcomes

Mohammad Saad Shaheed [a,1], Konstantina Gkritza [b,*]

[a] Institute for Transportation, Iowa State University, 2711 South Loop Drive, Suite 4700, Ames, IA 50010, United States
[b] Lyles School of Civil Engineering, HAMP G167B, Purdue University, West Lafayette, IN 47907, United States

ARTICLE INFO

Article history:
Received 19 September 2013
Received in revised form
12 March 2014
Accepted 12 March 2014

Keywords:
Crash severity
Heterogeneity
Latent class analysis
Motorcycle safety
Single-vehicle crashes

ABSTRACT

Unobserved heterogeneity has been recognized as a critical issue in traffic safety research that has not been completely addressed or often overlooked, and can lead to biased estimates and incorrect inferences if inappropriate methods are used. This paper uses a latent class approach to investigate the factors that affect crash severity outcomes in single-vehicle motorcycle crashes. Motorcycle crash data from 2001 to 2008 in Iowa were collected with a total of 3644 single-vehicle motorcycle crashes occurring during that time period. A latent class multinomial logit model is estimated that addresses unobserved heterogeneity by identifying two distinct crash data classes with homogeneous attributes. The estimation results show a significant relationship between severe crash injury outcomes and crash-specific factors (such as speeding, run-off road, collision with fixed object and overturn/rollover), riding on high-speed roads, riding on rural roads, riding on dry road surface, riding without a helmet, age (riders older than 25 years old) and impaired riding (riders under the influence of drug, alcohol or medication). The model fit and estimation results underline the need for segmentation of crashes, and suggest that the latent class approach can be a promising tool for modeling motorcycle crash severity outcomes.

1. Introduction

In the United States (U.S.), motorcycle registrations increased by 67% during the period from 2001 to 2010. However, motorcycles still constitute a small subset of the total registered motor vehicles, which makes up around 3% (National Highway Traffic Safety Administration, 2012). Despite their small share of vehicle registrations, motorcycles are over-represented in fatal crashes in the U.S. (Preusser et al., 1995). According to the National Highway Traffic Safety Administration (2012), motorcycle crash fatalities have been increasing each year in the U.S. from a historic low of 2116 fatalities in 1997 to as high as 5312 fatalities in 2008. In 2010, motorcyclists accounted for 14% of total traffic fatalities, although motorcycles accounted for only 0.6% of all vehicle miles traveled (National Highway Traffic Safety Administration, 2012). Similarly in the state of Iowa, motorcycle registrations increased by 41.4% from 2001 to 2010. Motorcycle fatalities in Iowa increased from 38 in 2001 to the highest of 61 in 2007, with motorcycle registrations increasing by 27% during that period.

* Corresponding author. Tel.: +1 765 494 4597.
E-mail addresses: mshaheed@iastate.edu (M.S. Shaheed), nadia@purdue.edu (K. Gkritza).
[1] Tel.: +1 515 708 5756.

http://dx.doi.org/10.1016/j.amar.2014.03.002
2213-6657/

Among the types of crashes involving motorcycles, single-vehicle motorcycle crashes have accounted for almost 50% of the total riders killed in both single- and multi-vehicle motorcycle crashes from 2001 to 2010 (National Highway Traffic Safety Administration, 2012). Previous research has asserted that single- and multi-vehicle crashes are best described when modeled separately (Geedipally and Lord, 2010).

Furthermore, some of the factors affecting the severity of motorcycle crashes are not observable, or relevant information is not reported by law enforcement agencies and cannot be collected from state crash records. Such information includes motorcycle speed, motorcycle size, riders' physical health, riders' experience and training. Therefore, heterogeneity may arise from these unobserved factors when developing motorcycle crash severity models as a function of observed factors, resulting in biased parameters and incorrect inferences (Mannering and Bhat, 2014). To overcome this problem, this paper uses a latent class approach to investigate the factors and their effect on single-vehicle motorcycle crash severity outcomes.

2. Literature review

Extensive research has been conducted on investigating the risk factors and patterns associated with single-vehicle motorcycle crashes. Mannering and Grodsky (1995) stated that since riding a motorcycle is generally recognized as a dangerous activity, it may tend to attract risk-seeking individuals, regardless of their age or their socioeconomic background. Risky driving behaviors, such as speeding or drinking and riding, can significantly affect the severity outcome of a crash. For instance, Shankar and Mannering (1996) found that speeding increases the likelihood of fatal or injury outcomes using single-vehicle motorcycle crash data in the state of Washington. Similar results are also reported in Savolainen and Mannering (2007) that examined crash data in Indiana. Past studies also have shown that motorcycle engine size is associated with motorcycle crash injury severity outcomes (de Lapparent, 2006; Harrison and Christie, 2005; Lin et al., 2003; Mullin et al., 2000; Pai and Saleh, 2007; Quddus et al., 2002; Savolainen and Mannering, 2007; Sexton et al., 2004).

Age and gender are among the most important rider-specific factors that affect both frequency and severity of motorcycle crashes. Previous research found that the likelihood of fatality and disabling injury in single-vehicle motorcycle crashes increases with increasing motorcycle rider age (Nunn, 2011; Pai and Saleh, 2007; Quddus et al., 2002; Savolainen and Mannering, 2007; Shankar and Mannering, 1996). The finding that older motorcyclists are more likely to be severely injured could be attributed to the slower reaction time and reduced sensory and perceptual ability, as well as to the decreased physical resiliency to motorcycle crashes compared to younger drivers (Pai and Saleh, 2007; Savolainen and Mannering, 2007).

Survey-based studies (Lin et al., 2003; Mullin et al., 2000) revealed that experience of the riders was an important risk factor associated with motorcycle crashes. Harrison and Christie (2005) found that riding patterns, skill acquisition, and training are important risk factors associated with motorcycle crashes. Other risk factors include a temporary break from riding and lack of basic training (Sexton et al., 2004). Note that these factors are usually not available in police-reported crash databases. Results of past research, as summarized in Lin and Krauss (2009), suggest that motorcycle riders are more likely to drink and ride compared to other motor-vehicle drivers. In the U.S., alcohol-related fatal motorcycle crashes are higher than alcohol-related fatal crashes involving other types of vehicles (National Highway Traffic Safety Administration, 2008).

The effectiveness of helmet use has been investigated extensively in the literature. Past work summarized in Lin and Krauss (2009), argued that helmet use reduced the risk of motorcycle deaths by 29% in the U.S. during the period 1972–1987. Based on the results of past studies in the U.S., it has been suggested that non-helmeted riders are more likely to die, incur head injuries, or require longer hospitalization compared to helmeted riders. Comprehensive helmet laws also have been associated with an increase in helmet use followed by a decrease in the number of motorcycle fatalities and head injuries (Houston, 2007; Lin and Krauss, 2009).

Road type and geometry along with roadside installations, pavement surface conditions, lighting and visibility conditions, and manner of crashes (such as run-off road, collision with stationary object or other) constitute another group of risk factors (Vlahogianni et al., 2012). With regards to lighting and visibility conditions, poor visibility due to horizontal or vertical curvature or darkness has been associated with an increase in motorcycle injury severity (Savolainen and Mannering, 2007). Turning to the type of collision, research has shown that collisions with stationary objects result in more severe injuries (Quddus et al., 2002; Savolainen and Mannering, 2007). Riding on wet roadway surface can be a risk factor as well; however, the crash risk may be mitigated, especially if motorcyclists maintain lower speed due to the poor surface condition (Savolainen and Mannering, 2007).

3. Methodological background

This section discusses the modeling techniques that have been employed in the past to estimate motorcycle crash severity outcomes. Shankar and Mannering (1996) applied a multinomial logit model to estimate the factors affecting single-vehicle motorcycle crash severity. Quddus et al. (2002) estimated ordered probit models to examine motorcycle damage and injury severity resulting from motorcycle crashes. Savolainen and Mannering (2007) estimated both nested logit and multinomial logit models to analyze single- and multi-vehicle motorcycle crashes in Indiana. Chimba and Sando (2010) applied multinomial logit and multinomial probit models to identify the factors associated with motorcycle crash injury severities in Florida. Rifaat et al. (2012) estimated an ordered logit model, a heterogeneous choice model, and a partially constrained generalized ordered logit model to identify factors contributing to increasing the severity of motorcycle crashes in Calgary, Canada. Cafiso et al. (2012) carried out a logistic regression analysis to identify the factors affecting the severity of crashes involving mopeds and motorcycles. Lately, Jones et al. (2013) estimated a multinomial logit regression model to examine the factors affecting motorcycle crash severity in Alabama using five years of crash data from 2006 to 2010.

Savolainen et al. (2011) discusses the limitations of discrete ordered (probit and logit models) and unordered models (such as the multinomial logit model) for studying crash injury severities. Traditional ordered probability models place restrictions on the way variables influence outcome probabilities. The proportional odds assumption in the ordered probability models allows for a positive coefficient to unambiguously indicate that the variable associated with the positive coefficient increases the likelihood of the highest ordered outcome occurring (such as fatality) and decreases the likelihood of the lowest ordered outcome (such as no injury).

On the other hand, each crash severity outcome has a different function in the multinomial logit (MNL) model making the model structure more flexible. However, the MNL model is particularly susceptible to correlation of unobserved effects from one crash-severity level to the next. Such correlation causes a violation of the model's independence of irrelevant alternatives (IIA) property. The random terms, ε_{ij}, in the crash severity functions in an MNL model are also assumed to be independent and identically distributed (IID). This assumption might be violated in practice as well, as crash severity functions do not contain an exhaustive list of all contributing factors.

Random parameters (mixed logit) and finite mixture (latent class) methods can accommodate individual unobserved heterogeneity by allowing parameters to differ across observations. In the case of a mixed logit model, the analyst needs to assume a distribution relating how parameters vary across observations or (group of observations) and/or determine observation groups. This requirement is relaxed in a latent class model, where a discrete distribution, represented by a finite and specified number of mass points, is used to identify homogeneous subgroups of data (Mannering and Bhat, 2014). A disadvantage of the latent class method is that it does not account for the possibility of variation within a class as it assumes homogeneous characteristics of the within-class observations (Mannering and Bhat, 2014).

A latent class approach has been frequently used for incorporating preference heterogeneity in marketing research (Bujosa et al., 2010; Greene and Hensher, 2003, 2013; Olaru et al., 2011; Shen, 2009; Wen and Lai, 2010; Wen et al., 2012) and lately, for analyzing driver injury severities (Chu, 2014; Eluru et al., 2012; Xie et al., 2012; Xiong and Mannering, 2013; Yasmin et al., 2014). Dataset-specific evidence in the literature suggest stronger statistical support for latent class logit models compared to conventional MNL models (Xie et al., 2012; Wen et al., 2012), ordered logit models (Chu, 2014), nested logit models (Wen et al., 2012), and mixed logit models (Greene and Hensher, 2003; Shen, 2009).

3.1. The latent class multinomial logit model

Single-vehicle motorcycle crash severity outcomes are estimated using a latent class multinomial logit model (LC-MNL). The LC-MNL model is considered a special form of the mixed (random parameter) multinomial logit model. The probability that an injury outcome j will occur in a typical mixed MNL model can be described as follows:

$$\text{Prob}\,(j) = \int \text{Prob}\,(j|\beta) f(\beta)\, d\beta \tag{1}$$

where Prob(j) is the probability of a crash to result in crash severity outcome j; β is a parameter vector; Prob($j|\beta$) is the conditional probability of a crash to result in crash severity outcome j given the estimated parameter vector β; and $f(\beta)$ is the density function of β.

The explicit specification of distribution of each random parameter in the vector β is challenging when applying a mixed MNL model. In the latent class formulation, parameter heterogeneity across observations is modeled with a discrete distribution or set of classes (Greene and Hensher, 2003; Greene, 2007). In this paper, it is assumed that the collected motorcycle crash dataset can be categorized into M different classes. Each crash event belongs to different classes with certain probabilities $Prob(class = m)$ that are specified by the MNL form and are not revealed to the analyst:

$$\text{Prob}\,(\text{class} = m) = \frac{\exp\,(\theta_m z_i)}{\sum_{m=1}^{M} \exp\,(\theta_m z_i)} \tag{2}$$

where θ_m are class-specific parameters and z is an optional set of crash-specific or other characteristics (motorcycle, rider attributes or other). The class specific probabilities maybe a set of fixed constants θ_m (the last of which is fixed at zero) if no z characteristics are observed. Note that there are no firm rules on the selection of the number of classes (Xie et al., 2012). In general, more classes would make model convergence more challenging and the interpretation of the model results more difficult (Eluru et al., 2012). It has been suggested to add one class at a time until further addition does not enhance intuitive interpretation and data fit (Eluru et al., 2012; Greene and Hensher, 2003). It is possible to compare a model with $M+1$ classes to one with M classes using likelihood ratio tests (Greene and Hensher, 2003), or based on the Bayesian Information Criterion (BIC) (Roeder et al., 1999; Boxall and Adamowicz, 2002). Most recent studies of crash injury severities have used the BIC measure to determine the number of classes (Chu, 2014; Yasmin et al., 2014). The BIC for a given empirical model is equal to

$$\text{BIC} = -2\text{LL} + K\,\ln\,(Q) \tag{3}$$

where LL is the log-likelihood value at convergence, K is the number of parameters, and Q is the number of observations. Lower BIC values indicate a better model fit.

Based on the above discussion, the unconditional probability that an injury outcome j will occur is given as

$$\text{Prob}\,(j) = \sum_{m=1}^{M} \text{Prob}\,(\text{class} = m) \times \text{Prob}\,(j|\text{class} = m) \tag{4}$$

where Prob(j|class $= m$) is the probability that the jth injury outcome may occur within a class and is assumed to be generated by the MNL model as follows:

$$\text{Prob}\,(j|\text{class} = m) = \frac{\exp\,(\beta_m X_{ij})}{\sum_{j=1}^{J} \exp\,(\beta_m X_{ij})} \tag{5}$$

where X_{ij} is a set of explanatory variables (crash-specific or other); β_m is a class-specific parameter vector that takes a finite set of values; and $j = 1, \ldots, J$ the number of crash severity outcomes. In this paper, three crash severity outcomes are considered: fatal and major injury, minor and possible/unknown injury, and no-injury crashes. Grouping some of the severities outcomes in one category was necessary to ensure adequate number of observations for modeling purposes.

3.2. Analysis of marginal effects

In a LC-MNL model, a marginal effect is calculated for each class using the same method as that of a conventional MNL model. For indicator variables, the marginal effects are computed as the difference in the estimated probabilities with the indicator variable changing from zero to one, while all the other variables are equal to their means. The direct and cross-marginal effects are calculated following Eqs. (6) and (7), respectively (Greene, 2007):

$$\frac{\partial P_{ij}}{\partial x_{ijk}} = \beta_{jk} P_{ij} (1 - P_{ij}) \tag{6}$$

$$\frac{\partial P_{iq}}{\partial x_{ijk}} = -\beta_{jk} P_{ij} P_{iq} \tag{7}$$

The direct marginal effect (Eq. (6)) represents the effect that a unit change in x_{ijk} has on the probability for crash i to result in outcome j (denoted by P_{ij}). The cross-marginal effect (Eq. (7)) shows the effect of a unit change in variable k of alternative j ($j \neq q$) on the probability (P_{iq}) for crash i to result in outcome q. Following Greene (2007) and Xie et al. (2012), the final marginal effect of a variable is calculated as the summation of the marginal effects for each class weighted by their posterior latent class probabilities. In cases where variables enter multiple severity functions, changing their values would affect all corresponding severity function values and probabilities. Their combined effects can be estimated by adding together the marginal effects with respect to different crash severity outcomes.

4. Data description

We collected data on single-vehicle motorcycle crashes that occurred from 2001 to 2008 in Iowa (a total of 3644 crashes). More than 50 variables were considered in the latent class analysis. Table 1 shows the summary statistics of the variables that were found to be significant in the LC-MNL model. The distribution of crashes by severity outcome shows that 90% of the single-vehicle crashes during that period resulted in an injury outcome. This highlights the importance of identifying the factors that are associated with each severity outcome and subsequently, proposing countermeasures to improve motorcycle safety.

Table 1 shows that speeding was reported as the major cause for 12% of the crashes and running-off the road was reported as the major cause for 25% of the crashes. Collisions with fixed objects and overturn/rollover were the most common harmful events. Crashes on high-speed facilities (speed limit greater than 55 mph) and on rural roads comprised 49% and 55% of the total number of single-vehicle motorcycle crashes, respectively. The majority of crashes occurred on dry roadway surface, and almost one third of crashes occurred under dark conditions. About one quarter of the riders involved in the crashes was less than 25 years old, while 10% of the riders were under the influence of drug, alcohol, or medication. Only 28% of the riders involved in single-vehicle motorcycle crashes from 2001 to 2008 wore helmets. The low helmet rate is likely a reflection of the fact that Iowa does not have a universal helmet law.

5. Estimation results

Two[2] distinct segments (classes) with homogeneous attributes were found significant; latent class 1 with probability 0.58 and latent class 2 with probability 0.42. The class specific probabilities are a set of fixed constants (see Eq. (2)),

[2] Latent class MNL models with both two and three classes were estimated and BIC values for each model were computed to identify the optimal number of latent classes The BIC values for the two-class and three-class models were 1490 (42 parameters) and 1540 (60 parameters), respectively. As such, a two-class MNL model was selected and presented in Table 3.

Table 1
Summary statistics of variables included in the LC-MNL model.

Variable mnemonic	Description	Mean (standard deviation)
SEVERITY	Fatal/major injury/minor injury/ possible or unknown/ no-injury	0.05/0.24/0.42/0.19/0.10
SPEED	Major cause: speeding (1=Yes, 0=No)	0.12 (0.32)
ROR	Major cause: run-off road (1=Yes, 0=No)	0.25 (0.43)
COLFIOB	1st harmful event: collision with fixed object (1=Yes, 0=No)	0.29 (0.45)
OVERTRN	1st harmful event: overturn/rollover (1=Yes, 0=No)	0.22 (0.41)
COLNFIX	1st harmful event: collision with non-fixed object (1=Yes, 0=No)	0.25 (0.43)
ECONANI	Contributing circumstances: animal on roadway (1=Yes, 0=No)	0.13 (0.33)
ROADSIDE	Location of 1st harmful event: roadside (1=Yes, 0=No)	0.07 (0.26)
SPLIG55	Speed limit: greater than 55 mph (1=Yes, 0=No)	0.49 (0.50)
RURAL	Crash on a rural road (1=Yes, 0=No)	0.55 (0.49)
DRY	Surface condition: dry (1=Yes, 0=No)	0.88 (0.33)
DARK	Light condition: dark (1=Yes, 0=No)	0.31 (0.47)
WEEKEND	Crash on a weekend (1=Yes, 0=No)	0.43 (0.49)
AUG	Month: August (1=Yes, 0=No)	0.16 (0.37)
SUMM	Month of crash: May, June, or July (1=Yes, 0=No)	0.48 (0.49)
MALE	Rider gender: male (1=Yes, 0=No)	0.92 (0.27)
YOUNG	Rider age: less than 25 years old (1=Yes, 0=No)	0.25 (0.43)
DRALC	Influence of drug, medicine or alcohol (1=Yes, 0=No)	0.10 (0.31)
HELMET	Helmet used by rider (1=Yes, 0=No)	0.28 (0.45)

Table 2
Characteristics of LC-MNL model with two classes.

Components	Classes	
	Class 1	Class 2
Class characteristics		
Crash population share	0.58	0.42
Crash injury severity		
No-injury	0.037	0.036
Minor/possible injury	0.718	0.542
Fatal/major injury	0.245	0.422

as examining segmentation on the basis of crash-specific characteristics did not result in a superior model in terms of data fit. The crash severity outcome shares within each class are shown in Table 2. It is shown that a crash allocated to class 2 is more likely to result in a more severe injury outcome than if allocated to class 1.

Table 3 shows the latent class multinomial logit estimation results. A total of 19 variables were found to be statistically significant, mostly at a 0.05 significance level, and the McFadden Pseudo R-Square was 0.334. These variables can be grouped into crash-specific characteristics, location and time of crash characteristics, roadway and environmental conditions as well as rider attributes. The LC-MNL results for each class show that each variable has two sets of parameters associated with it. Note that some of the parameters have the same sign across the two classes (for example, speeding, riding on dry surface condition and helmet use), while others have opposite signs (for example, influence of drug, medicine or alcohol, riding in summer months) or are not significant in both classes (for example, run-off road, collision with fixed object, overturn/rollover, age), suggesting that there is heterogeneity between the two classes. As such, the interpretation of the model results cannot be based on the magnitude and sign of the parameters but rather it is based on marginal effects.

Table 4 shows the direct and cross marginal effects as well as the combined marginal effects for the variables entering multiple severity functions. For example, the variable named COLNFIX was present in the severity functions of both minor/possible injury outcome and no-injury outcome. The combined marginal effect of the variable of COLNFIX for fatal/major injury outcome was calculated by adding the cross marginal effects for minor/possible injury outcome (1.990) no-injury outcome (0.453) resulting in a value of 2.443. Since the original marginal effects are very small, the estimated marginal effects were multiplied by 100, expressing percentage-point differences.

Speeding as the major cause of a crash was associated with fatal/major injury outcomes; this finding is line with past research (Savolainen and Mannering, 2007; Shankar and Mannering, 1996). The probability of a fatal/major crash would increase by 2.7% when the rider of a motorcycle is reported to either drive too fast for the conditions or exceed authorized speed limit. A similar effect was found for crashes with run-off road reported as their major cause.

The analysis results also show that the probability of a fatal/major injury outcome increased by 2.9% in the case of a collision between a motorcycle and a fixed object. This is expected as fixed objects are not likely to absorb much energy during a crash causing severe injuries to motorcycle riders. Previous studies also showed that collision with stationary

Table 3
LC-MNL model estimation results.

Variable mnemonic	**Latent class 1**						**Latent class 2**					
	Fatal/major injury		**Minor/possible injury**		**No-injury**		**Fatal/major injury**		**Minor/possible injury**		**No-injury**	
	Parameter	*t*-Statistic	Parameter	*t*-Statistic	Parameter	*t*-Statistic	Parameter	*t*-Statistic	Parameter	*t*-Statistic	Parameter	*t*-Statistic
SPEED	1.289	3.28	–	–	–	–	1.429	5.168	–	–	–	–
ROR	−0.459	−1.129	–	–	–	–	2.243	8.747	–	–	–	–
COLFIOB	1.139	2.925	–	–	–	–	0.156	0.608	–	–	–	–
OVERTRN	−0.084	−0.201	–	–	–	–	0.988	3.711	–	–	–	–
COLNFIX	–	–	−0.09	−0.223	−0.019	−0.035	–	–	−1.285	−4.881	−2.511	−4.517
ECONANI	–	–	–	–	−0.356	−0.573	–	–	–	–	1.857	3.792
ROADSIDE	1.509	2.961	–	–	–	–	0.594	1.63	–	–	–	–
SPLIG55	–	–	0.689	1.617	0.689	1.617	–	–	−2.289	−8.432	−2.289	−8.432
RURAL	0.721	1.868	–	–	–	–	−0.554	−2.076	–	–	–	–
DRY	0.895	2.433	2.656	5.326	–	–	1.048	3.092	−2.274	−4.516	–	–
DARK	–	–	–	–	0.503	1.327	–	–	–	–	0.025	0.078
WEEKEND	–	–	–	–	−1.412	−2.985	–	–	–	–	0.489	1.591
AUG	−0.842	−2.236	–	–	–	–	0.781	2.912	–	–	–	–
SUMM	–	–	2.528	5.047	–	–	–	–	−1.864	−9.036	–	–
MALE	–	–	−0.686	−1.544	–	–	–	–	5.933	12.149	–	–
YOUNG	−1.377	−3.510	–	–	–	–	−0.010	−0.048	–	–	–	–
HELMET	−0.052	−2.236	–	–	–	–	−1.134	−5.468	–	–	–	–
DRALC	3.374	5.302	–	–	–	–	−1.396	−4.631	–	–	–	–
Latent Class Probability (t-statistic)					0.58 (33.983)				0.42 (13.886)			
Number of Observations								3644				
Log-likelihood at zero								−2007.165				
Log-likelihood at convergence								−1336.326				
McFadden Pseudo-R^2								0.334				
BIC								1490				

Table 4
Estimated marginal effects of the variables included in the LC-MNL model.

Variable mnemonic	In severity function of	Effects on probabilities of the severity outcomes (%)		
		Fatal/major	Minor/possible	No-injury
SPEED	Fatal/major injury	2.681	−2.454	−0.165
ROR	Fatal/major injury	2.028	−1.863	−0.165
COLFIOB	Fatal/major injury	2.911	−2.634	−0.277
OVERTRN	Fatal/major injury	1.095	−0.961	−0.134
	Minor/possible injury	1.99	−2.16	0.17
COLNFIX	No-injury	0.453	0.255	−0.708
	Combined effects	2.443	−1.905	−0.538
ECONANI	No-injury	−0.359	−0.187	0.546
ROADSIDE	Fatal/major injury	1.223	−1.129	−0.094
	Minor/possible injury	3.916	−3.738	−0.178
SPLIG55	No-injury	0.615	−0.1781	−0.437
	Combined effects	4.531	−3.916	−0.615
RURAL	Fatal/major injury	0.923	−0.874	−0.048
	Fatal/major injury	10.468	−9.426	−1.043
DRY	Minor/possible injury	−2.885	4.951	−2.066
	Combined effects	7.583	−4.475	−3.109
DARK	No-injury	−0.15	−0.328	0.478
WEEKEND	No-injury	0.011	0.257	−0.269
AUG	Fatal/major injury	−0.209	0.103	0.106
SUMM	Minor/possible injury	−0.299	0.240	−0.058
MALE	Minor/possible injury	−25.724	28.929	−3.205
YOUNG	Fatal/major injury	−1.44	1.254	0.186
HELMET	Fatal/major injury	−1.702	1.442	0.260
DRALC	Fatal/major injury	1.859	−1.802	−0.056

objects result in more severe injuries (Keng, 2005; Lin et al., 2003; Quddus et al., 2002; Savolainen and Mannering, 2007). Crashes involving overturned motorcycles had a 1.1% higher likelihood to result in a fatal or major injury outcome. Another interesting finding from our analysis results is the lower likelihood of minor/possible injuries for motorcycles colliding with non-fixed objects (such as debris on roadways, work zone equipment, or other). It seems that single-vehicle motorcycle crashes with motorcycles colliding with fixed or non-fixed objects were less likely to result in minor/possible outcome and more likely to result in a fatal/major injury outcome.

Motorcyclists were less likely to sustain injuries in crashes where animals on the roadway were reported as a contributing circumstance to the crashes. Single-vehicle motorcycle crashes had a 1.2% higher likelihood to result in severe injury outcomes when the first harmful event occurred on the roadside. This finding is also in line with previous research (Savolainen and Mannering, 2007) showing that the outcome of motorcycles colliding with roadside objects such as curbs or tree poles is more likely to be severe. Single-vehicle motorcycle crashes occurring on roads with speed limit higher than 55 mph were more likely to result in fatal or major injury outcomes and less likely to result in minor/possible or no-injury outcomes. This may be attributed to higher speeds on these roads. Single-vehicle motorcycle crashes occurring on rural roads resulted in a 0.9% higher probability of a fatal or major injury.

Motorcyclists faced higher risk of sustaining a fatal/major injury in a single-vehicle crash on a dry pavement surface, possibly because of risk compensating behavior, as argued in past work (Shaheed et al., 2013). Crashes occurring during dark conditions were less likely to be fatal, which might seem counterintuitive. This might be attributed to the possibility that motorcyclists are generally more cautious when riding in the dark. Single-vehicle motorcycle crashes occurring on a weekend were more likely to result in severe and minor/possible injury outcomes. Crashes occurring during the heart of the riding season (May through July) were more likely to result in minor/possible injury outcomes and less likely to result in severe injury outcomes. This finding suggests that motorcyclists are more likely to be severely injured in a crash occurring at the beginning of the riding season, probably because of the time elapsed since they last rode a motorcycle. A similar crash severity outcome is also likely towards the end of the riding season, probably because of changes in motorcyclists' risk perception as their exposure increases and they become more confident of their riding skills. Moreover, motorcyclist protective clothing and helmets, which can decrease the probability of severe motorcycle crash severity outcomes, are less tolerable in warm weather (Shaheed et al., 2013).

Turning to the motorcycle rider characteristics, a male rider involved in a single-vehicle motorcycle crash was more likely to be less severely injured compared to a female rider. This result might be a reflection of the difference in driving style, experience, risk perception, and behavior between male and female riders (Savolainen and Mannering, 2007). Motorcyclists younger than 25 years old were less likely to be severely injured in single-vehicle motorcycle crashes compared to older motorcyclists, which is in line with previous research (Nunn, 2011; Pai and Saleh, 2007; Savolainen and Mannering, 2007; Shankar and Mannering, 1996). This in an important finding as demographic information on motorcycle riders shows a dramatic increase in older riders in recent years.

Results reveal that single-vehicle motorcycle crashes with helmeted riders were less likely to result in a fatal or major injury outcome, and more likely to result in a minor, possible/unknown or no-injury outcome. Previous research has shown that helmet use is positively associated with minor injury outcomes for single-vehicle motorcycle crashes (Savolainen and Mannering, 2007), as well as that interactions of no-use of helmet and other factors increase the likelihood of fatality and evident injuries in single-vehicle motorcycle crashes (Shankar and Mannering, 1996). This is an important finding considering the fact that there is no helmet law in Iowa.

Lastly, single-vehicle motorcycle crashes with riders under the influence of drug, medication or alcohol were almost 2% more likely to result in fatal or major injury outcomes. This finding is consistent with previous studies (Shankar and Mannering, 1996) reporting that alcohol-impaired riding increases the likelihood of fatality in single-vehicle motorcycle crashes. This finding also underscores the potential of alcohol, drug or medicine to increase the severity of head injuries once a crash has occurred (Waller et al., 1986).

6. Conclusions

This paper estimated a latent class multinomial logit model of single-vehicle motorcycle crash severities to correct for unobserved heterogeneity as an alternative to the mixed logit (random parameters) model. Motorcycle crash data from 2001 to 2008 in Iowa were collected with a total of 3644 single-vehicle motorcycle crashes occurring during that time period. Three crash injury outcomes were considered in this paper by combining fatal and major crash injury outcomes in one category, minor and possible crash injury outcomes in a second category, and no-injury crash outcomes in a third category. Two distinct subgroups of crashes with homogeneous attributes were identified with specified probabilities, and a total of 19 variables were found to be statistically significant, mostly at a 0.05 significance level. The interpretation of the estimation results was based on marginal effects.

The estimation results showed a significant relationship between severe crash injury outcomes and factors such as speeding, run-off road, collision with fixed object, overturn/rollover, first harmful event taking place on roadside, riding on high-speed roads, riding on rural roads, riding on dry road surface, riders older than 25 years old, riding without a helmet, and riders under the influence of drug, alcohol or medication. In view of our results, recommendations can be made to make motorcycle riders aware of the beneficial effect of helmet use and the adverse effects of impaired riding and speeding on motorcycle safety.

While our results are generally in line with past studies that examined the factors affecting single-vehicle motorcycle crash severity (mostly by estimating MNL models), these studies overlooked the issue of unobserved heterogeneity. We found that the estimated parameters of some of the significant variables had the same sign across the two classes (such as speeding, riding on dry road surface and helmet use), while others had opposite signs (such as impaired riding, riding on rural road) or were not significant in both classes (such as run-off road, collision with fixed object, overturn/rollover, age, and others). The fact that the impact of exogenous factors differed in sign, magnitude, and significance between the two classes underlines the need for segmentation of crashes. The estimation results and model fit suggest that the latent class approach can be a promising tool for modeling motorcycle crash severity outcomes as an alternative to a mixed logit model that requires specifying the probability distributions of each random parameter as well as a greater computational effort. However, more empirical studies are needed to apply this method to datasets in other states with different demographics, climate, helmet laws, and crash reporting systems in order to fully assess the effect of the factors affecting single-vehicle motorcycle crash severity as well as fully understand the strengths and weaknesses of the latent class approach.

Acknowledgment

The authors thank Mr. Dimitrios Bilionis for his assistance with data collection and the anonymous reviewers for their useful feedback.

References

Boxall, C., Adamowicz, W., 2002. Understanding heterogeneous preferences in random utility models: a latent class approach. Environmental and Resource Economics 23 (4), 421–446.

Bujosa, A., Riera, A., Hicks, R., 2010. Combining discrete and continuous representation of preference heterogeneity: a latent class approach. Environmental and Resource Economics 47 (4), 477–493.

Cafiso, S., La Cava, G., Pappalardo, G., 2012. A logistic model for powered two-wheelers in Italy. Procedia – Social and Behavioral Sciences 53, 881–890.

Chimba, D., Sando, T., 2010. Multinomial probability assessment of motorcycle injury severities. Advances in Transportation Studies 21, 73–80.

Chu, H.C., 2014. Assessing factors causing severe injuries in crashes of high-deck buses in long-distance driving on freeways. Accident Analysis and Prevention 62, 130–136.

de Lapparent, M., 2006. Empirical Bayesian analysis of accident severity for motorcyclists in large French urban areas. Accident Analysis and Prevention 38 (2), 260–268.

Eluru, N., Bagherib, M., Miranda-Moreno, L., Fu, L., 2012. A latent class modeling approach for identifying vehicle driver injury severity factors at highway–railway crossings. Accident Analysis and Prevention 47, 119–127.

Geedipally, S.R., Lord, D., 2010. Investigating the effect of modelling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson-gamma models. Accident Analysis and Prevention 42 (4), 1273–1282.

Greene, W.H., 2007. NLOGIT Version 4.0 Reference Guide. Econometric Software, Inc., Plainview, NY.

Greene, W.H., Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B 37 (8), 681–698.
Greene, W.H., Hensher, D.A., 2013. Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics 45 (14), 1897–1902.
Harrison, W.A., Christie, R., 2005. Exposure survey of motorcyclists in New South Wales. Accident Analysis and Prevention 37 (3), 441–451.
Houston, D.J., 2007. Are helmet laws protecting young motorcyclists? Journal of Safety Research 38 (3), 329–336.
Jones, S., Gurupackiam, S, Walsh, J., 2013. Factors influencing the severity of crashes caused by motorcyclists: analysis of data from Alabama. Journal of Transportation Engineering 139 (9), 949–956.
Keng, S.H., 2005. Helmet use and motorcycle fatalities in Taiwan. Accident Analysis and Prevention 37 (2), 349–355.
Lin, M.R., Chang, S.H., Pai, L., Keyl, P.M., 2003. A longitudinal study of risk factors for motorcycle crashes among junior college students in Taiwan. Accident Analysis and Prevention 35 (2), 243–252.
Lin, M.-R., Kraus, J.F., 2009. A review of risk factors and patterns of motorcycle injuries. Accident Analysis and Prevention 41 (4), 710–722.
Mannering, F.L., Grodsky, L.L., 1995. Statistical analysis of motorcyclists' perceived accident risk. Accident Analysis and Prevention 27 (1), 21–31.
Mannering, F.L., Bhat, C.R., 2014. Analytic methods in accident research: Methodological frontier and future directions. Analytic Methods in Accident Research 1, 1–22.
Mullin, B., Jackson, R., Langley, J., Norton, R., 2000. Increasing age and experience: are both protective against motorcycle injury? A case control study. Injury Prevention 6 (1), 32–35.
National Highway Traffic Safety Administration, 2008. Traffic Safety Figures. Center for Statistics and Analysis, Washington, DC ⟨http://www.nrd.nhtsa.dot.gov/Pubs/811155.PDF⟩.
National Highway Traffic Safety Administration, 2012. Traffic Safety Facts, Center for Statistics and Analysis, Washington, DC ⟨http://www.nrd.nhtsa.dot.gov/Pubs/811639.PDF⟩.
Nunn, S., 2011. Death by motorcycle: background, behavioral, and situational correlates of fatal motorcycle collisions. Journal of Forensic Sciences 56 (2), 429–437.
Olaru, D., Smith, B., Taplin, J., 2011. Residential location and transit-oriented development in a new rail corridor. Transportation Research Part A 45 (3), 219–237.
Pai, C.-W., Saleh, W., 2007. An analysis of motorcyclist injury severity under various traffic control measures at three-legged junctions in the UK. Safety Science 45 (8), 832–847.
Preusser, D.F., Williams, A.F., Ulmer, R.G., 1995. Analysis of fatal motorcycle crashes: crash typing. Accident Analysis and Prevention 27 (6), 845–851.
Quddus, M.A., Noland, R.B., Chin, H.C., 2002. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. Journal of Safety Research 33 (4), 445–462.
Rifaat, S.M., Tay, R., Barros, A., 2012. Severity of motorcycle crashes in Calgary. Accident Analysis and Prevention 49 (1), 44–49.
Roeder, K., Lynch, K., Nagin, D., 1999. Modeling uncertainty in latent class membership: a case study in criminology. Journal of the American Statistical Association 94, 766–776.
Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accident Analysis and Prevention 39 (5), 955–963.
Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention 43 (3), 1666–1676.
Sexton, B., Baughan, C., Elliott, M., Maycock, G., 2004. The Accident Risk of Motorcyclists. TRL Report TRL607, UK.
Shaheed, M., Gkritza, K., Zhang, W., Hans, Z., 2013. A mixed logit analysis of two-vehicle crash severities involving a motorcycle. Accident Analysis and Prevention, Special Issue on Emerging Research Methods and their Application to Road Safety 61, 119–128.
Shankar, V., Mannering, F., 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. Journal of Safety Research 27 (3), 183–194.
Shen, J., 2009. Latent class model or mixed logit model? A comparison by transport mode choice data. Applied Economics 41 (22), 2915–2924.
Vlahogianni, E.I., Yannis, G., Golias, J.C., 2012. Overview of critical risk factors in powered-two-wheeler safety. Accident Analysis and Prevention 49 (1), 12–22.
Waller, P. F, Stewart, J.R., Hansen, A. R, Stutts, J., Popkin, C.L., Rodgman., E., 1986. The potentiating effects of alcohol on driver injury. The Journal of the American Medical Association 256 (11), 1461–1466.
Wen, C.-H., Lai, S.-C., 2010. Latent class models of international air carrier choice. Transportation Research Part E 46 (2), 211–221.
Wen, C.-H., Wang, W.-C., Fu, C., 2012. Latent class nested logit model for analyzing high-speed rail access mode choice. Transportation Research Part E 48 (2), 545–554.
Xie, Y., Zhao, K., Huynh, N., 2012. Analysis of driver injury severity in rural single-vehicle crashes. Accident Analysis and Prevention 47, 36–44.
Xiong, Y., Mannering, F.L., 2013. The heterogeneous effects of guardian supervision on adolescent driver-injury severities: a finite-mixture random-parameters approach. Transportation Research Part B 49 (1), 39–54.
Yasmin, S., Eluru, N., Bhat, C.R., Tay, R., 2014. A latent segmentation based generalized ordered logit model to examine factors influencing driver injury severity. Analytic Methods in Accident Research 1, 23–38.