Accident Analysis and Prevention 101 (2017) 124–134

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Modeling single-vehicle run-off-road crash severity in rural areas: Accounting for unobserved heterogeneity and age difference

Linfeng Gong, Wei (David) Fan*

Department of Civil and Environmental Engineering, University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223, USA

ARTICLE INFO

Article history:
Received 18 October 2016
Received in revised form 8 January 2017
Accepted 13 February 2017

Keywords:
Roadway safety
Crash-injury severities
Run-off-road crashes
Age differences
Mixed logit model

ABSTRACT

This study investigates factors that significantly contribute to the severity of driver injuries resulting from single-vehicle run-off-road (SV ROR) crashes. A mixed logit model approach is employed to explore the potential unobserved heterogeneous effects associated with each age group: young (ages 16–24), middle-aged (ages 25–65), and older drivers (ages over 65). Likelihood ratio tests indicated that the development of separate injury severity models for each age group is statistically superior to estimating a single model using all data. Based on the crash data collected from 2009 to 2013 in North Carolina, a series of driver, vehicle, roadway, and environmental characteristics are examined. Both parameter estimates and their elasticities are developed and used to interpret the models. The estimation results show that contributing factors which significantly affect the injury severity of an SV ROR crash differ across three age groups. Use of restraint device and horizontal curves are found to affect crash injuries and fatalities in all age groups. Reckless driving, speeding, distraction, inexperience, drug or alcohol involvement, presence of passengers, and driving an SUV or a van are found to have a more pronounced influence in young and middle-aged drivers than older drivers. Compared to the passenger cars, older drivers are less likely to experience possible injuries in a large-size vehicle (e.g., truck or bus). The average annual daily traffic volume and lighting conditions are also found to influence the resulting injury severity of an SV ROR crash specific to young drivers.

1. Introduction

Run-off-road (ROR) crashes occurring on rural highways account for the majority of highway fatalities. According to the statistical data released by the FHWA's Roadway Departure Safety Program (Federal Highway Administration, 2014), there were 17,791 people killed as a result of roadway departure crashes, which was 54% of the traffic fatalities in the U. S. in 2014. Compared to other crash types, ROR crashes are typically associated with more severe crash consequences when the vehicle left the roadway and hit fixed roadside objects; or when the vehicle crossed the centerline and was involved in a head-on crash. As emphasized in the FHWA's Roadway Departure Safety Program (Federal Highway Administration, 2014), overturns, crossing a center line or median, and involving trees or shrubs on the roadside account for more than 70% of all ROR crashes. Run-off-road crashes are more likely to occur on rural roads rather than in urban environments. In a previous study conducted by Liu and Subramanian (2009), it was found that ROR crashes accounted for 80.6% and 56.2% of all crashes on rural and urban roadways, respectively. This is probably due to the fact that the types of driving that occur in rural areas are inherently different from those in urban environments. On rural highways, speeds are likely to be higher and fatigue may be more prevalent. In addition, less roadway safety countermeasures are implemented on rural highways due to their relatively lower traffic volumes. Therefore, developing an effective approach to investigating the unique characteristics and contributing factors associated with single-vehicle run-off-road (SV ROR) crashes can provide valuable insights into how to decrease traffic injuries in rural areas.

A number of studies have been undertaken to investigate causal factors that may affect the occurrence and the resulting injury severity outcome of ROR crashes (e.g., Roque and Cardoso, 2014; Peng and Boyle, 2012; Liu and Ye, 2011). The present study focuses primarily on identifying factors contributing to SV ROR crashes and their influence on injury severity. Based on the discrete nature of crash severity data (e.g., fatal, injury, and property damage only (PDO)), several discrete choice models have been applied to investigate factors that may affect the outcome level of SV ROR crashes.

* Corresponding author.
E-mail addresses: lgong@uncc.edu (L. Gong), wfan7@uncc.edu, wei.d.fan@gmail.com (W. Fan).

http://dx.doi.org/10.1016/j.aap.2017.02.014

For example, Liu and Ye (2011) utilized a binary logit model to evaluate the impacts of both driver- and vehicle-related attributes on ROR injury severity. Palamara et al. (2013) developed a multinomial logit (MNL) model to identify factors affecting SV ROR crashes in Western Australia. It is noteworthy that there are a number of factors that may affect the resulting injury severity of an SV ROR crash, including vehicle characteristics, roadway features, traffic-related factors, environmental conditions, and the complex interactions between them. Many of the existing crash severity models are developed on the basis of the police-reported data, which contains only a portion of the explanatory variables. However, some of the many factors that affect the resulting injury severity of a crash still remain unknown to the analyst. Such factors constitute the unobserved heterogeneity across individual injury observations and, if not accounted for, may result in biased parameter estimates and erroneous statistical inferences (McFadden and Train, 2000; Train, 2009; Mannering et al., 2016). As an example, one can consider age as an observed explanatory variable in modeling crash injury severity outcomes. While there are apparent physiological and psychological differences across various age groups (e.g., young versus older drivers), the reactions times, risk-taking behaviors, and physiological characteristics can vary significantly from one individual to another, even for those within the same age group (e.g., two young drivers may have distinct driving behaviors). In most cases, a driver's physiological and psychological features, and their driving behaviors are generally unknown to the analyst; age is just a proxy for these unobserved factors. Mannering et al. (2016) pointed out that the heterogeneous effects of age may be even more pronounced than other explanatory variables as it is sometimes included as a grouped indicator variable in modeling crash outcome levels (Cerwick et al., 2014; Behnood and Mannering, 2016).

This paper intends to contribute to a better understanding of the factors that affect SV ROR crashes occurring on rural highways by employing a mixed logit model approach to accounting for the unobserved heterogeneous effects across the population. Meanwhile, the present study also investigates the similarities and dissimilarities among factors that significantly contribute to the injury severity of SV ROR crashes among different age groups. The results of this study offer valuable insights into the underlying relationship between risk factors and SV ROR injury severity and thus help promote the implementation of more effective countermeasures to mitigate SV ROR crash severity in rural areas.

2. Literature review

2.1. Previous researches related to SV ROR crash severity modeling

A number of discrete choice models have been applied to explain the injury severity outcomes of ROR crashes (Peng and Boyle, 2012; Liu and Ye, 2011; Schneider et al., 2009; Palamara et al., 2013; Hu and Donnell, 2010; Lu et al., 2010; Deng et al., 2006). For example, Peng and Boyle (2012) previously defined crash outcome as a binary variable (injury/fatal versus PDO) and used a binary logit model to explore the impact of commercial driver factors on crash consequence levels of SV ROR crashes. The results indicated that speeding, drowsiness and fatigue, distraction, and inattention negatively affected the outcome severity. Schneider et al. (2009) developed separate MNL models to assess driver injury severity resulting from single-vehicle crashes along various horizontal curves (small, medium, and large radius), respectively. In this study, crash injury outcomes were split into 5 categories: fatal, incapacitating injury, non-incapacitating injury, possible injury and PDO. Their results showed that injuries were more likely on curves with a moderate radius of between 500 and 2800 ft than on the other two types of curves. Palamara et al. (2013) developed an MNL model to identify factors affecting SV ROR crashes in Western Australia. Their findings highlighted the role of speed, road alignment, and type of collision (hit object or others) as contributors to injury severity. Hu and Donnell (2010) estimated a nested logit model of median barrier crash severities using a five-year crash dataset collected on rural highways in North Carolina. The model estimates showed that increasing the median barrier offset was associated with a lower probability of severe crash outcomes.

To account for the ordinal nature of crash severity data, traditional ordered probability models have been applied in modeling ROR crash injury levels as well. Lu et al. (2010) employed the traditional ordered logit (ORL) and probit (ORP) models to examine the nexus between crash severity and safety-related factors for cross-median crashes that occurred from 2001 to 2007 in Wisconsin. Both models found that more severe injuries occurred on roadways posted with higher speed limits. Deng et al. (2006) developed an ORP model to analyze the statistical association between head-on crash severity and potential causal factors, such as the geometric characteristics of the road segment, weather conditions, road surface conditions, and time of occurrence.

2.2. Statistical analyses of crash-injury severities

Although the MNL model and the traditional ordered response models have been widely used in crash severity modeling, there are some inherent limitations that may restrict the application of those structures. Specifically, the MNL model assumes that the random terms of each severity function is independent and identically distributed (IID) from another. In reality, however, this is not always the case due to potential correlations between unobserved factors (Abdel-Aty, 2003; Xie et al., 2012). A famous example of the violation of the IID assumption is the red bus/blue bus problem. In terms of the traditional ordered response models (i.e., ORL and ORP), they hypothesized that the impacts of various exogenous factors on crash injury severity are constant across all severity levels. This is referred to as the proportional odds or parallel regression assumption in literature (Williams, 2006; Savolainen et al., 2011).

In recent years, a myriad of advanced econometric frameworks have been applied in injury severity analyses, including the generalized ordered logit (GOL) model (Eluru, 2013), partial proportional odds (PPO) model (Wang and Abdel-Aty, 2008; Quddus et al., 2010), latent-class (finite mixture) model (Xie et al., 2012; Behnood et al., 2014; Behnood and Mannering, 2016), and mixed logit model (Moore et al., 2011; Kim et al., 2013; Behnood and Mannering, 2016). A more comprehensive review of crash injury severity models can found in (Savolainen et al., 2011; Mannering and Bhat, 2014). These advanced model frameworks enable researchers to breakthrough the limitations of the MNL and the traditional ordered probability models. For example, the GOL model relaxes the parallel assumption of the traditional ordered probability models by allowing model parameters to vary freely across severity levels. Eluru (2013) compared the performance of the GOL model with the traditional ORL model and the MNL model, finding that the GOL model consistently outperformed the other two. As the inherent parallel assumption in traditional ordered response models may not necessary hold over all severity levels, the PPO model relaxes the parallel assumption by allowing a portion of the explanatory variables to affect each level of the response variable differently, while other independent predictors can still adhere to the parallel assumption. Examples of using PPO models in crash severity modeling can be found in Wang and Abdel-Aty (2008) and Quddus et al. (2010).

As for unordered discrete outcome structures, the latent-class and mixed logit models have been proposed for modeling crash injury severity levels to address the potential limitations of the independence from irrelevant alternatives (IIA) property associ-