Accident Analysis and Prevention 96 (2016) 79–87

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Examining driver injury severity outcomes in rural non-interstate roadway crashes using a hierarchical ordered logit model

Cong Chen[a], Guohui Zhang[a,*], Helai Huang[b], Jiangfeng Wang[c], Rafiqul A. Tarefder[d]

[a] Department of Civil and Environmental Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, HI 96822, United States
[b] Urban Transport Research Center, School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan 410075, China
[c] School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, China
[d] Department of Civil Engineering, University of New Mexico, Albuquerque, NM 87131, United States

ARTICLE INFO

Article history:
Received 26 August 2015
Received in revised form 13 February 2016
Accepted 20 June 2016

Keywords:
Driver injury severity
Rural non-interstate roadway
Hierarchical ordered logit model
Full Bayesian inference
Traffic safety

ABSTRACT

Rural non-interstate crashes induce a significant amount of severe injuries and fatalities. Examination of such injury patterns and the associated contributing factors is of practical importance. Taking into account the ordinal nature of injury severity levels and the hierarchical feature of crash data, this study employs a hierarchical ordered logit model to examine the significant factors in predicting driver injury severities in rural non-interstate crashes based on two-year New Mexico crash records. Bayesian inference is utilized in model estimation procedure and 95% Bayesian Credible Interval (BCI) is applied to testing variable significance. An ordinary ordered logit model omitting the between-crash variance effect is evaluated as well for model performance comparison. Results indicate that the model employed in this study outperforms ordinary ordered logit model in model fit and parameter estimation. Variables regarding crash features, environment conditions, and driver and vehicle characteristics are found to have significant influence on the predictions of driver injury severities in rural non-interstate crashes. Factors such as road segments far from intersection, wet road surface condition, collision with animals, heavy vehicle drivers, male drivers and driver seatbelt used tend to induce less severe driver injury outcomes than the factors such as multiple-vehicle crashes, severe vehicle damage in a crash, motorcyclists, females, senior drivers, driver with alcohol or drug impairment, and other major collision types. Research limitations regarding crash data and model assumptions are also discussed. Overall, this research provides reasonable results and insight in developing effective road safety measures for crash injury severity reduction and prevention.

1. Introduction

Although less frequent in occurrence compared to urban crashes, rural crashes result in more severe injuries and deaths due to the relatively high traffic speed and considerable portion of heavy vehicles in rural areas. According to the National Highway Traffic Safety Adminisitration (NHTSA) (2013), the national fatality rate is 1.37 per 100 million vehicle miles traveled (VMT) for rural areas, which is almost twice as high as the urban fatality rate (0.7 per 100 VMT). A consistent crash fatality pattern is also demonstrated at the regional level. In 2011, 351 fatalities occurred from traffic crashes in New Mexico, and 241 of these crashes occurred at rural locations, even though rural crashes only accounted for 17.6% of all reported crashes (NMDOT, 2012). Among these rural fatalities, 73.9% happened at rural non-interstate locations, despite rural interstate highways carrying the primary portion of rural traffic volume. Therefore, the significant life lost and emotional burden call for an in-depth investigation of traffic crash risk factors at rural non-interstate locations in order to reduce crash injury severities. Existing studies regarding rural traffic safety issues focus on rural highways, including interstate and non-interstate highways (Cafiso et al., 2010; Park et al., 2012; Viner, 1995), and specific locations, such as intersections (Kim et al., 2007; Laberge et al., 2006; Ye et al., 2009). However, traffic crashes occurring on rural roadways are seldom included in these studies; therefore rural crash characteristics were not comprehensively examined. Considering the severe crash outcomes at rural non-interstate locations in New Mexico, this study aims to examine the injury severity patterns at all of

* Corresponding author.
E-mail addresses: cchen02@unm.edu (C. Chen), guohui@hawaii.edu (G. Zhang), huanghelai@csu.edu.cn (H. Huang), wangjiangfeng@bjtu.edu.cn (J. Wang), tarefder@unm.edu (R.A. Tarefder).

http://dx.doi.org/10.1016/j.aap.2016.06.015
0001-4575/

these locations, including rural intersections, rural local roadways and rural non-interstate highways.

Among all the techniques applied in traffic safety analyses, generalized regression models are the major type utilized to formulate the nature of crash occurrences and injury outcomes. Existing models consider injury severity as a binary-response variable, such as no injury versus injury or no/slight injury versus severe injury/death, and binary logit or probit models are typically applied. Due to the ordinal nature of crash injury severities, ordinal-response models have been increasingly used in traffic safety analyses (Bhat and Eluru, 2009; Bhat et al., 2014; Eluru and Bhat, 2007; Eluru, 2013; Eluru et al., 2008; Li et al., 2012; Yasmin et al., 2015, 2014a,b; Ye and Lord, 2014). For instance, Yasmin et al. (2014a,b) utilized a generalized ordered logit model with latent segmentation variables to investigate heterogeneous factors and their respective impacts on driver injury severity outcomes. With a mixed generalized ordered-response model, Eluru et al. (2008) assessed the injury severity patterns of pedestrians and bicyclists in traffic crashes. They concluded that individual age, roadway speed limit, crash location and crash time are the most important factors related to these non-motorist injury severity outcomes.

Traffic crash data are generally organized in a hierarchical structure, as illustrated by Huang and Abdel-Aty (2010). Hierarchical models are able to capture these structural features expressed as between-crash or between-segment variance and outperform generalized logit or probit models. Different from conventional frequency-based estimation methods, Bayesian inference method is able to utilize extra knowledge regarding the research topic besides the studied dataset, such as historical patterns regarding a target type of crashes, as the informative prior for the posterior estimation on the parameters of interest, and is increasing its popularity in traffic safety studies (Yu and Abdel-Aty, 2013). Moreover, due to the limited availability of accurate prior information, Bayesian non-informative priors are widely accepted in a significant amount of studies (Deublein et al., 2013; Haque et al., 2010; Huang et al., 2008; MacNab, 2003; Xie et al., 2013; Yu and Abdel-Aty, 2014), and it provided at least compatible performance in model fit and parameter estimation, even though these settings compromises the merit of Bayesian inference method to some extent. For example, Haque et al. (2010) investigated the risk factors for motorcycle-related crashes at intersections using hierarchical Bayesian Poisson models. Deublein et al. (2013) presented a hierarchical Bayesian approach for road accident prediction by incorporating gamma distribution, multivariate Poisson-lognormal regression and Bayesian inference together.

The significant loss in rural non-interstate traffic crashes justifies the need for a comprehensive examination of the injury severity patterns and the risk factors related to injury severity outcomes in these crashes. Ordered-response model design is able to capture the ordinal nature of crash injury severities, and hierarchical model structure is capable of simulating the variance among different crashes as well as capturing the correlations among vehicles in the same crash. This paper applies a hierarchical ordered logit model to examine the contributing factors on driver injury severity outcomes in rural non-interstate crashes based on two-year crash data in New Mexico from 2010 to 2011. According to the standard police report, driver injury severity was recorded as a five-level outcome: no injury, compliant of injury/possible injury, visible injury, incapacitating injury and death. An FB inference method was implemented for parameter estimation, and Bayesian non-informative prior was used for parameter posterior estimation. The significant attributes were retained and redundant ones were removed based on 95% BCI. An ordinary ordered logit model was utilized to analyze the same dataset for model performance comparison purposes. The rest of this paper is organized as follows: the next section presents a comprehensive and in-depth literature review regarding rural crashes, ordered-response models, and hierarchical Bayesian approaches. Detailed descriptions and pre-processing of the studied dataset are illustrated in Section 3. Section 4 introduces the implemented model design and specifications. Research results are explicitly discussed in Section 5, followed by Section 6 summarizing the limitations of this study. The research is then concluded in Section 7.

2. Literature review

Although it is less populated in rural areas, traffic crashes occurring in rural locations result in more severe injuries and fatalities. As a result, tremendous efforts have been made to examine the characteristics of these crashes. Karlaftis and Golias (2002) discovered that rural traffic volume and roadway geometric characteristics are two important risk factors for rural crashes. Polus and Cohen (2012) proposed a Poisson model to predict accident frequency on rural roadways with low traffic volumes, which demonstrated a comparable fit with negative binomial modeling. Brown and Cline (2001) explored contributing factors of injury severity for rear-seat occupants in rural traffic crashes, discovering that intoxicated driving and senior age would lead to more severe injuries. Primary studies regarding rural crashes focus on crash occurrences and severity outcomes on rural highways, including interstate and non-interstate highways. Abbas (2004) evaluated safety performance of rural roadways and developed several statistical models with different functional forms for crash frequency and fatality prediction. A considerable amount of rural highway safety studies focus on crash patterns on rural two-lane highways. Cafiso et al. (2010) developed a safety performance function (SPF) model for rural two-lane highways considering factors regarding traffic exposure, roadway geometry, environmental elements, etc. Park et al. (2012) verified the effect of wider edge lines on rural two-lane highways in improving traffic safety and approved their universal implementation. Numerous studies have also addressed traffic crash features on rural interstate highways. Sun and Garber (2002) applied an empirical Bayes (EB) method in exploring the safety effects of rural interstate speed. Wagenaar et al. (1990) conducted time-series analyses to estimate the impact of a 65 mph speed limit on rural interstate highways on mortality and injury. Viner (1995) found that 46% of rural interstate fatal and severe injury crashes resulted from rollovers, of which 35% are on their left sides. Other studies also shed light on rural intersections. Laberge et al. (2006) proposed the framework of an intersection decision support (IDS) system to extract traffic information in order to facilitate drivers' car-following strategies at rural stop-controlled intersections. Kim et al. (2007) investigated traffic crash risk factors for different severities at rural intersections through binomial hierarchical multilevel models. Ye et al. (2009) developed multi-dimensional Poisson models to examine the potentials of different types of crashes at rural intersections. However, no previous research has been conducted to comprehensively investigate driver injury patterns and related attributes for non-interstate rural locations, which motivates the authors to meet this gap.

Numerous studies have been conducted using generalized regression models with ordered responses to capture the ordinal nature of crash injury severity and the influence of risk factors on crash occurrence and injury outcomes. Yamamoto and Shankar (2004) developed a bivariate ordered-response probit model to evaluate driver and passenger injury severity patterns in fix-object collision accidents regarding driver characteristics, vehicle object features, natural conditions, etc. Haleem and Abdel-Aty (2010) examined crash injury severities at intersections without signalized control using an ordered probit model with a 5-injury-level response and identified the contributory factors among traffic con-