Accident Analysis and Prevention 97 (2016) 69–78

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Driver injury severity outcome analysis in rural interstate highway crashes: a two-level Bayesian logistic regression interpretation

Cong Chen[a], Guohui Zhang[a,*], Xiaoyue Cathy Liu[b], Yusheng Ci[c], Helai Huang[d], Jianming Ma[e], Yanyan Chen[f], Hongzhi Guan[g]

[a] Department of Civil and Environmental Engineering, University of Hawaii at Manoa, 2540 Dole Street, Honolulu, HI 96822, United States
[b] Department of Civil & Environmental Engineering, University of Utah, 110 Central Campus Drive, Suite 2000, Salt Lake City, UT 84112, United States
[c] School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China
[d] Urban Transport Research Center, School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan 410075, China
[e] Traffic Operations Division, Texas Department of Transportation, Austin, TX, 78717, USA
[f] Beijing Transportation Engineering Key Laboratory, Beijing University of Technology, Beijing, 100124, China
[g] Transportation Research Center, Beijing University of Technology, Beijing, 100124, China

ARTICLE INFO

Article history:
Received 22 February 2016
Received in revised form 7 June 2016
Accepted 22 July 2016

Keywords:
Driver injury severity
Rural interstate highway
Traffic crash
Hierarchical model
Bayesian inference

ABSTRACT

There is a high potential of severe injury outcomes in traffic crashes on rural interstate highways due to the significant amount of high speed traffic on these corridors. Hierarchical Bayesian models are capable of incorporating between-crash variance and within-crash correlations into traffic crash data analysis and are increasingly utilized in traffic crash severity analysis. This paper applies a hierarchical Bayesian logistic model to examine the significant factors at crash and vehicle/driver levels and their heterogeneous impacts on driver injury severity in rural interstate highway crashes. Analysis results indicate that the majority of the total variance is induced by the between-crash variance, showing the appropriateness of the utilized hierarchical modeling approach. Three crash-level variables and six vehicle/driver-level variables are found significant in predicting driver injury severities: road curve, maximum vehicle damage in a crash, number of vehicles in a crash, wet road surface, vehicle type, driver age, driver gender, driver seatbelt use and driver alcohol or drug involvement. Among these variables, road curve, functional and disabled vehicle damage in crash, single-vehicle crashes, female drivers, senior drivers, motorcycles and driver alcohol or drug involvement tend to increase the odds of drivers being incapably injured or killed in rural interstate crashes, while wet road surface, male drivers and driver seatbelt use are more likely to decrease the probability of severe driver injuries. The developed methodology and estimation results provide insightful understanding of the internal mechanism of rural interstate crashes and beneficial references for developing effective countermeasures for rural interstate crash prevention.

1. Introduction

Compared to urbanized areas, rural areas have a higher potential of inducing more severe driver injuries in traffic crashes in spite of a lower crash frequency (Eiksund, 2009; Jones et al., 2008). According to the National Highway Traffic Safety Administration (NHTSA) (NHTSA, 2013), 54% of total fatal crashes and 55% of total fatalities occurred in US rural areas. This is particularly striking because only 19% of the total population lives in these areas. At the regional level in New Mexico, there were 350 fatalities due to traffic accidents in 2011, and 273 of them were on rural roadways (NMDOT, 2012). Averaged by travel distance, 1.37 fatalities per 100 million vehicle miles traveled (VMT) occur in rural areas, which is much higher than the 0.7 fatalities per 100 million VMT occurring in urban areas (U.S. Department of Transportation, 2013). Therefore, significant research efforts are needed to investigate the internal mechanism of rural crashes in order to develop effective countermeasures to reduce rural crash frequency and casualties.

Rural highways are major corridors carrying a significant portion of high speed traffic and are prone to inducing traffic accidents with severe injuries. Tremendous effort has been made to address traffic safety issues on rural highways. Governmental agencies and researchers have been investigating the impacts of rural interstate speed limits on traffic casualties at national and regional levels since the 1980s (Cannon et al., 2009; Federal Highway Administration (FHWA), 2004; McCarthy, 1993). With the develop-

* Corresponding author.
E-mail address: guohui@hawaii.edu (G. Zhang).

http://dx.doi.org/10.1016/j.aap.2016.07.031
0001-4575/

ment of mathematical models and computing techniques in recent years, significant research has been conducted to address rural highway safety issues by examining factors regarding infrastructure, vehicles, human behavior and traffic environments and their effects on crash frequency and injury severity (Cafiso et al., 2010; de Oña et al., 2013; Lord et al., 2005; Siskind et al., 2011). For example, using a multivariate analysis, Siskind et al. (2011) found that speeding, alcohol involvement and traffic rule violations were major factors resulting in fatal crashes in rural areas. Lord et al. (2005) proposed predictive models to investigate the association between crash-flow-density and crash-flow-volume/capacity ratio on rural freeway sections and assessed their influence on crash frequencies.

Compared with the maximum likelihood estimation (MLE) method, Bayesian inference method is able to produce more reliable results; therefore it has been increasingly associated with traffic safety studies in recent years (Elvik, 2013; Haque et al., 2010; Huang and Abdel-Aty, 2010; Yanmaz-Tuzel and Ozbay, 2010; Yu and Abdel-Aty, 2013). For instance, Elvik (2013) evaluated the influence of speed limits on traffic crashes using the empirical Bayes (EB) method. Yanmaz-Tuzel and Ozbay (2010) utilized full Bayes (FB) models to explore the most effective countermeasures in reducing crash frequencies. Hierarchical Bayesian methods allow for the examination of different levels of information, such as crash and vehicle levels. They also produce estimation results with Bayesian inference techniques, making them preferable for contemporary transportation researchers. Several studies have been conducted to prove the superiority of hierarchical Bayesian models in traffic safety analysis. Yu and Abdel-Aty (2014) applied a hierarchical Bayesian binary probit model into investigating significant factors on injury severity in highway crashes and compared it with binary probit models with maximum likelihood estimation. They concluded that the proposed hierarchical model is superior in examining crash injuries with more explanatory attributes. Xie et al. (2013) examined the effectiveness of a proposed Bayesian hierarchical negative binomial model, a regular Bayesian negative binomial model and a Bayesian random effect model in predicting signalized-intersection crash frequency and found that the Bayesian hierarchical negative binomial model outperforms the others. Other peer studies employed hierarchical Bayesian models to explore the significant factors of crash frequencies or severities in different scenarios (Abdalla, 2005; Ahmed et al., 2011; Chen et al., 2016a; Deublein et al., 2013; Huang et al., 2008).

These studies provide a comprehensive review of rural traffic safety analyses and an in-depth understanding of hierarchical Bayesian models in addressing traffic safety issues. However, there were no studies conducted to analyze driver injury severities in crashes happening on rural interstate highways using the hierarchical Bayesian method. Thus, the authors were motivated to conduct this research. In this paper, a hierarchical binary logistic regression model was developed to investigate the significant factors and their respective effects on driver injury severities in rural interstate crashes based on crash data collected in the State of New Mexico from 2010 to 2015. Driver injury severities were modified as a binary outcome: low injury severity levels, including no injury and compliant of injury (coded as 0) and high injury severity levels, including visible injury, incapacitating injury and fatality (coded as 1). The explanatory variables were classified as two levels of hierarchy, crash-level variables and driver/vehicle-level variables, revealing information regarding crash information, weather and environment information, vehicle factors, driver behavior and demographic characteristics, etc. Significant attributes were selected based on a 95% Bayesian credible interval (BCI). The rest of this paper is organized as follows: a comprehensive literature review regarding rural road safety analysis and hierarchical Bayesian applications is provided in Section 2. Data descriptions are illustrated in Section 3, followed by the description of the proposed methodology in Section 4. The model estimation results are explicitly discussed in Section 5, and research limitations are summarized in Section 6. This paper is concluded in Section 7.

2. The state of the art and the practice

Significant research has been done to address traffic safety issues on rural highways with different models from various aspects. Khorashadi et al. (2005) explored the difference in driver injuries between rural and urban highway crashes with truck involvement through a multinomial logit (MNL) model. Farah et al. (2009) developed a Tobit model to examine the relationship between crash potential and drivers' passing behavior on rural two-lane highways and concluded that driver characteristics are among the significant attributes affecting crash potential. Cafiso et al. (2010) developed synthetic analysis models to evaluate the safety performance of two-lane rural highways based on crash risk factors regarding exposure geometry, consistency and context information. Using a mixed logit model, Chen and Chen (2011) evaluated the distinctiveness in the injury patterns of truck drivers in single-vehicle and multi-vehicle crashes on rural highways and assessed their risk factors. None-regression and non-parametric models were also utilized in traffic safety studies. Karlaftis and Golias (2002) employed a hierarchical tree-based regression model in rural highway accident analysis. Kashani and Mohaymany (2011) applied classification tree models to predict injury severity patterns of two-lane rural roadway traffic accidents. de Oña et al. (2013) examined the primary factors contributing to rural highway crash severity through latent class clustering (LCC) and Bayesian network (BN) techniques and concluded that the synthetic use of these techniques outperforms separate applications.

Rural interstate highways play an indispensable role in the U.S. national highway system by carrying massive long-distance and high-speed transportation. Rural interstate safety issues have been in the spotlight regarding speed limit settings. Gallaher et al. (1989) discovered that a speed limit increase from 55 mph to 65 mph on rural interstate highways led to an increase of 2.9 fatal crashes per 100 million VMT. McCarthy (1993) discovered that the rise of rural interstate speed limits significantly increases the occurrence of alcohol-related crashes regardless of crash types. In a nationwide survey, the NHTSA (1998) found that an increase of rural interstate speed limits from 55 mph to 65 mph results in about 30% more fatalities than the expected values. FHWA (2004) investigated the influences of uniform speed limits and discrepant speed limits for passenger cars and trucks on the safety performance of rural interstate highways. They concluded that vehicle mean speed and crash frequency are inclined to increase in both scenarios. Cannon et al. (2009) developed an automated method based on naturalistic driving data extracted from traffic videos and data streams to analyze rural road crashes, including rural interstate highways. However, there is no previous research conducted to comprehensively examine the contributing factors to rural interstate crashes.

Bayesian inference methods and derivative models have been more associated with contemporary traffic safety analyses. de Lapparent (2006) employed an EB model to evaluate the impact of socio-demographic elements on motorcycle accident severity. Persaud et al. (2010) examined the effectiveness of EB and FB models in before-after road safety assessment. Eksler (2010) investigated local-level roadway safety performance in both spatial and temporal domains via FB model. Hierarchical Bayesian models synthesize the advantages of Bayesian inference method and hierarchical models and have become increasingly preferable in traffic safety studies. MacNab (2003) proposed a Bayesian hierarchical Poisson regression model to facilitate accident monitoring and prevention in both spatial and temporal domains. Abdalla