Transportation Research Part A 75 (2015) 30–41

Contents lists available at ScienceDirect

Transportation Research Part A

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

Assessment of the effects of highway geometric design features on the frequency of truck involved crashes using bivariate regression

Chunjiao Dong [a,*], Shashi S. Nambisan [b], Stephen H. Richards [a], Zhuanglin Ma [c]

[a] *Center for Transportation Research, The University of Tennessee, 600 Henley Street, Knoxville 37996, TN, USA*
[b] *Department of Civil & Environmental Engineering, The University of Tennessee, 319 John D. Tickle Building, Knoxville, TN 37996-2321, USA*
[c] *School of Automobile, Chang'an University, Nan Er Huan Zhong Duan, Xi'an 710064, Shaanxi, China*

A R T I C L E I N F O

Article history:
Received 28 July 2014
Received in revised form 4 February 2015
Accepted 5 March 2015
Available online 28 March 2015

Keywords:
Large truck involved crashes
Crash frequency
Geometric design features
Bivariate ZINB model
Bivariate NB model

A B S T R A C T

Given the enormous losses to society resulting from large truck involved crashes, a comprehensive understanding of the effects of highway geometric design features on the frequency of truck involved crashes is needed. To better predict the occurrence probabilities of large truck involved crashes and gain direction for policies and countermeasures aimed at reducing the crash frequencies, it is essential to examine truck involved crashes categorized by collision vehicle types, since passenger cars and large trucks differ in dimensions, size, weight, and operating characteristics. A data set that includes a total of 1310 highway segments with 1787 truck involved crashes for a 4-year period, from 2004 to 2007 in Tennessee is employed to examine the effects that geometric design features and other relevant attributes have on the crash frequency. Since truck involved crash counts have many zeros (often 60–90% of all values) with small sample means and two established categories, car-truck and truck-only crashes, are not independent in nature, the zero-inflated negative binomial (ZINB) models are developed under the bivariate regression framework to simultaneously address the above mentioned issues. In addition, the bivariate negative binomial (BNB) and two individual univariate ZINB models are estimated for model validation. Goodness of fit of the investigated models is evaluated using AIC, SBC statistics, the number of identified significant variables, and graphs of observed versus expected crash frequencies. The bivariate ZINB (BZINB) models have been found to have desirable distributional property to describe the relationship between the large truck involved crashes and geometric design features in terms of better goodness of fit, more precise parameter estimates, more identified significant factors, and improved predictive accuracy. The results of BZINB models indicate that the following factors are significantly related to the likelihood of truck involved crash occurrences: large truck annual average daily traffic (AADT), segment length, degree of horizontal curvature, terrain type, land use, median type, lane width, right side shoulder width, lighting condition, rutting depth (RD), and posted speed limits. Apart from that, passenger car AADT, lane number, and indicator for different speed limits are found to have statistical significant effects on the occurrences of car-truck crashes and international roughness index (IRI) is significant for the predictions of truck-only crashes.

* Corresponding author.

http://dx.doi.org/10.1016/j.tra.2015.03.007

1. Introduction

The number of total registered trucks in the United States has grown steadily during the 10-year period, from 94.9 million in 2003 to about 133.1 million in 2012, a 40% increase (FHWA, 2003–2012). Large trucks have made an important contribution to the growth of the national economy by facilitating the distribution of a large portion of the nation's products [more than 13 billion tons per year (USDOT/BTS, 2013)]. However, as truck transport has become more important, the safety issues associated with the large truck traffic, in terms of traffic crashes, injuries, and fatalities have become more evident. Indeed, large trucks are a serious safety concern. More than 3000 people are killed every year in large truck involved crashes. The yearly number of fatalities in large truck involved crashes is on the rise again after nearly a decade of declining. According to the most recent data from the National Highway Traffic Safety Administration, in 2012, 3921 people were killed and 104,000 people were injured in crashes involving large trucks and 83% of those killed and 76% of those injured were not occupants of the trucks (NHTSA, 2012). A significant proportion of deaths occur to passenger car occupants because the large sizes and heavy weights of large trucks transmit most of the crash force to the passenger cars.

Truck issues currently command a large share of the traffic safety interest of researchers, policy makers, and the public. In recent years, some studies have been performed to better identify and understand factors that contribute to the frequencies of large truck involved crashes. Geometric design features, environmental factors, traffic conditions, and driver performances have been investigated to determine how these factors influence the occurrences of truck crashes. However, most of the research focused on modeling truck crash counts as a whole. Given the differences observed in the characteristics associated with car-truck crashes and truck-only crashes, some researchers (Dong et al., 2014) have proposed using separate crash prediction models for these crashes. Separate predictive models have been developed to estimate the safety performances when, instead of total crash counts, one wishes to model specific types of crash counts, such as single-vehicle crashes versus multi-vehicle crashes (Geedipally and Lord, 2010), and crashes resulting in injuries and non-injuries (Khattak et al., 2002). These efforts indicated that developing two distinct models provide better predicting performances than developing models that combine both crash categories together. Though the separate models have been employed as a starting point for modeling the counts of specific types of crashes (as opposed to total crashes), researchers (Geedipally and Lord, 2010; Dong et al., 2014) have found that crash counts exhibit characteristics that make the application of the separate models problematic. Specifically, separate models cannot handle the correlation among two interdependence crash types. In this paper, the relationship between the geometric design features, pavement conditions, traffic factors, and crash frequencies are investigated by using bivariate regression models. Since truck involved crash counts have a significant amount of zeros, the NB and ZINB regressions are developed under the bivariate regression framework for jointly modeling car-truck and truck-only crashes simultaneously.

2. Literature review

Most of the recent research on improving truck traffic safety have focused on reducing the degree of injury sustained by those involved in truck crashes and a wide variety of methodological techniques have been applied to gain a thorough understanding of the factors that affect the degree of injury, given that a crash has occurred (Lyman and Braver, 2003; Chen and Chen, 2011; Lemp et al., 2011; Chang and Chien, 2013; Zhu and Srinivasan, 2011). These studies provide valuable insights into the safety of trucks. However, it is important to note that reducing crash frequencies and reducing crash-injury severities require different strategic approaches. With regards to the crash frequency study, the relationships between total crashes and geometric design features of road segments, such as horizontal curvature, vertical grade, lane width, and shoulder width have been studied in numerous previous studies [a thorough review and assessment of these studies can be found in a recent paper (Lord and Mannering, 2010)]. Compared to the predictive models developed for total crashes, the studies that investigated the effects of highway geometric design features on the occurrences of truck involved crashes are relatively limited. Over the past 20 years, only a few researchers have proposed predictive models specifically to analyze the frequency of crashes involving large trucks.

Blower et al. (1993) analyzed crash rates of heavy truck-tractors using log-linear methods. They investigated the effects of the number of trailers, road type, area type, and time of day on the rates of casualty and property-damage-only crashes. The results showed that differences between tractors with one and two trailers were not significant and characteristics of the operating environment were found to have significant effects on crash rates. In other words, the results showed that operating environment is more important in determining the risk of crash involvements than truck configuration. This study provided considerable interest both in its method and its results. However, log-linear models only demonstrate association between variables and the investigated variables are all treated as "response variables". If one or more variables are treated as explicitly dependent and others as independent, in case of crash frequency analysis, then the log-linear models are not appropriate and the results obtained from these models are questionable. Subsequent studies have used more flexible methods and have achieved slightly different conclusions from Blower et al. (1993). A case-control study was performed to address whether multiple-trailer combination vehicles were overinvolved in crashes on interstate highways in Indiana by Braver et al. (1997). The day of week, time of day, urban/rural area, and specific highway were identified as significant factors using logistic regression model. The results indicated that the crash risk of double-trailer trucks might be greater under less favorable operating conditions and road environments compared to the single-trailer trucks. The major limitation of this