Accident Analysis and Prevention 62 (2014) 87–94

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Differences in passenger car and large truck involved crash frequencies at urban signalized intersections: An exploratory analysis

Chunjiao Dong[a,*], David B. Clarke[a], Stephen H. Richards[b], Baoshan Huang[b]

[a] Center for Transportation Research, The University of Tennessee, 600 Henley Street, Knoxville, TN 37996, USA
[b] Department of Civil & Environmental Engineering, The University of Tennessee, 72 Perkins Hall, Knoxville, TN 37996, USA

ARTICLE INFO

Article history:
Received 3 June 2013
Received in revised form 21 August 2013
Accepted 17 September 2013

Keywords:
Crash frequency
Geometric design
MVPLN model
Bayesian method

ABSTRACT

The influence of intersection features on safety has been examined extensively because intersections experience a relatively large proportion of motor vehicle conflicts and crashes. Although there are distinct differences between passenger cars and large trucks-size, operating characteristics, dimensions, and weight-modeling crash counts across vehicle types is rarely addressed. This paper develops and presents a multivariate regression model of crash frequencies by collision vehicle type using crash data for urban signalized intersections in Tennessee. In addition, the performance of univariate Poisson-lognormal (UVPLN), multivariate Poisson (MVP), and multivariate Poisson-lognormal (MVPLN) regression models in establishing the relationship between crashes, traffic factors, and geometric design of roadway intersections is investigated. Bayesian methods are used to estimate the unknown parameters of these models. The evaluation results suggest that the MVPLN model possesses most of the desirable statistical properties in developing the relationships. Compared to the UVPLN and MVP models, the MVPLN model better identifies significant factors and predicts crash frequencies. The findings suggest that traffic volume, truck percentage, lighting condition, and intersection angle significantly affect intersection safety. Important differences in car, car–truck, and truck crash frequencies with respect to various risk factors were found to exist between models. The paper provides some new or more comprehensive observations that have not been covered in previous studies.

1. Introduction

Traffic crashes at urban intersections place a huge burden on society in terms of death, injury, lost productivity, and property damage. Based on the fatality analysis reporting system (FARS) and national automotive sampling system-general estimates system (NASS-GES) data, about 40% of the estimated 5,338,000 crashes during 2011 in the United States were intersection-related. Of those intersection crashes, about 36 percent occurred at signalized intersections. Furthermore, signalized intersections also tended to experience more severe crashes. Injury crashes accounted for 33.2% of reported signalized intersection crashes, compared to 25.2% for non-signalized intersection crashes.

Turns with obstructed view, illegal maneuvers, false assumptions about other drivers' actions, and misjudgment of gap or conflicting vehicle speed are usually considered as the main contributors to intersection crashes. Non-standard geometric features and mixed traffic conditions increase the potential for intersection crashes. Intersection design considers the largest vehicle likely to use the facility with considerable frequency; this may present difficulties for vehicles with different sizes and dimensions. In addition, some vehicle (e.g., large commercial trucks) weights and dimensions have increased substantially, with important implications for highway design criteria, especially intersections. This raises safety questions, such as whether older intersection designs are adequate to serve these larger trucks and passenger cars simultaneously and, specifically, which geometric conditions pose the most serious safety problems for large trucks and passenger cars.

Passenger cars and large trucks share the same intersection driving environment, but differ in size, dimensions, weight, and operating characteristics. Therefore, to provide useful guidelines for crash prevention, it is critically important to examine crash frequencies categorized by vehicle type. Using multivariate regression, this study investigates the relationship between intersection geometric features, traffic factors, and crash frequencies. Our paper has several objectives. We identify factors contributing to intersection-related crashes for crash counts by vehicle type. We predict crash frequencies for intersections with specified attributes. Given a set of geometric and traffic factors, we identify which variables are relatively more critical to the safety performance of the intersection. Finally, we estimate the percentage reduction in

* Corresponding author. Tel.: +1 8659741826.
E-mail address: CDONG5@UTK.EDU (C. Dong).

0001-4575/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aap.2013.09.011

crashes resulting from various improvements in intersection geometry.

2. Literature review

Models to estimate crash frequency on roadway segments or at intersections fall into two broad categories. One category includes conventional univariate regression models, such as the Poisson model, Poisson-gamma (negative binomial) model, Poisson-lognormal model, zero-inflated model, Conway–Maxwell–Poisson model, gamma model, and generalized estimating equation model. The second category includes potentially more realistic specifications such as generalized additive models, random-effects models, negative multinomial models, random-parameters models, bivariate/multivariate models, finite mixture/Markov switching models, duration models, and hierarchical/multilevel models (for a complete review of this literature see Lord and Mannering, 2010).

The literature pertaining specifically to intersection crash counts since 2010 is sparse. We category this research into three classes based on its objectives. The first research class addresses the crash count as a whole and focuses on developing more sophisticated methodology. A second class focuses on the effects of new technology installed at intersections or some typical geometric features on crash frequencies. A third class focuses on analyzing frequencies of a specific crash type.

Several papers have described recent attention to analytical method development. Castro et al. (2012) proposed an equivalent latent variable-based generalized ordered response framework for count data models. Using this proposed framework, they gainfully introduced temporal and spatial dependencies through the latent continuous variables. Their formulation also allows handling excess zeros in correlation count data. Qin and Reyes (2011) developed an alternative modeling approach-quantile regression-to handle the heterogeneity issue in crash count data and to demonstrate the resulting disparity in crash trends. Mitra and Washington (2012) used random-effects negative binomial and random-parameters negative binomial models incorporating spatial variables in intersection crash research. The researchers listed here focused on overcoming issues related to combining data from different locations and times to increase sample size. The methodological improvements include considering spatial and/or temporal effects and considering random-parameters instead of fixed-parameters. However, the research may be too limited to draw any sound conclusions.

Donnell et al. (2010) used a negative binomial model to explore the safety effects of roadway lighting at intersections. Hu et al. (2011) used a Poisson model to estimate the effects of red light camera enforcement on fatal crashes. Wu et al. (2013) developed a random-parameters negative binomial model to evaluate safety impacts of signal-warning flashers and speed control at high-speed signalized intersections. Those researchers are more focused on method application than methodological innovation. In other words, their results address the effects of using certain devices or technologies. The model approach is based on the researchers' data and prior experience, so these papers are more useful as application examples.

In the third class, researchers used crash data to develop models to explain the relationship between frequency and various explanatory factors for a particular crash type. Since crash count data are likely to be correlated, the researchers have to deal with this issue when modeling crash counts. Haque et al. (2010) investigated motorcycle crashes at four-legged and T signalized intersections and used random-effects Poisson models to address correlation and heterogeneity issues. The result shows that the model allowing autoregressive lag-1 dependence specification in the error term is the most suitable one (random-effects Poisson (AR-1) model). Pulugurtha and Sambhara (2011) developed several binomial negative models to estimate pedestrian crashes, but didn't report the correlation issue. From these papers, random-effects models seem appropriate when research focuses on a particular crash type, but fail when one wants to investigate the relationships between the attribute factors and several crash types.

Based on literature since 2010, one may conclude that intersection crash frequency research is limited and the results inconclusive. Since our study investigated the relationship between geometric features, traffic factors, and crash frequencies by collision vehicle type, it seemed natural to consider a multivariate model. Recently, multivariate regression models have become popular in the traffic safety field, especially for crash rate/frequency estimation. Ma and Kockelman (2006) used an MVP regression model to estimate the injury count by severity level. Positive correlation in unobserved factors affecting count outcomes was found across severity levels, resulting in a statistically significant assistive latent term. Several papers (Park and Lord, 2007; Ma et al., 2008; EI-Basyouny and Sayed, 2009) described a MVPLN model approach to explore the relationship between factors and crash counts. Anastasopoulos et al. (2012) proposed a multivariate tobit regression model to handle left-censored at zero issues. These efforts demonstrate the appropriateness of multivariate models in crash-frequency modeling when, instead of total crash counts, one wishes to model specific types of crash counts.

Although multivariate regression models already have been developed in statistics, there have been almost no attempts to employ those models in intersection safety to model data on multivariate crash frequencies. This paper develops and presents a multivariate regression model of crash frequencies by collision vehicle type. We use crash data for signalized intersections in Tennessee. In addition, the paper investigates the performance of UVPLN, MVP, and MVPLN regression models in establishing the relationship between crashes and geometric design of roadway intersections.

3. Model structure and estimation

Given that crash-frequency is a non-negative integer, most recent research has used the Poisson regression model as a starting point. However, researchers often find that crash data characteristics make using the simple Poisson regression problematic. Specifically, Poisson models cannot handle over- and under-dispersion and they can be adversely affected by low sample-means and can produce biased results in small samples (Lord and Mannering, 2010). The UVPLN regression model is an extension of the Poisson model to overcome possible over-dispersion in the data. However, neither Poisson nor UVPLN models account for correlations of different collision types. Crash counts (across vehicle types) effectively require a multivariate model because the response is multi-dimensional. In this study, a MVPLN model is taken for frequency estimation. The following section presents the general forms of the MVPLN regression model and provides brief descriptions of its estimation procedures.

3.1. MVPLN model

For a set of data on road collisions at n intersections, where the collisions at each intersection are classified into K categories, define the vector $\mathbf{Y}_i = (y_{i1}, y_{i2}, \ldots, y_{iK})'$, where y_{ij} denote the number of collisions at the ith location in category j. Suppose that y_{ij} follows a Poisson distribution with mean $\lambda_{ij} = \exp(x_i\beta_j + \varepsilon_{ij})$, that is

$$y_{ij}|\varepsilon_i, \beta_j \sim \text{Poisson} \quad (\lambda_{ij}) \tag{1}$$