Accident Analysis and Prevention 111 (2018) 12–22

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Analysis of crash proportion by vehicle type at traffic analysis zone level: A mixed fractional split multinomial logit modeling approach with spatial effects

Jaeyoung Lee*, Shamsunnahar Yasmin, Naveen Eluru, Mohamed Abdel-Aty, Qing Cai

Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, FL 32816-2450, United States

ARTICLE INFO

Keywords:
Multinomial logit fractional split model
Traffic crash analysis
Macroscopic crash analysis
Traffic analysis zones
Vehicle type
Screening

ABSTRACT

In traffic safety literature, crash frequency variables are analyzed using univariate count models or multivariate count models. In this study, we propose an alternative approach to modeling multiple crash frequency dependent variables. Instead of modeling the frequency of crashes we propose to analyze the proportion of crashes by vehicle type. A flexible mixed multinomial logit fractional split model is employed for analyzing the proportions of crashes by vehicle type at the macro-level. In this model, the proportion allocated to an alternative is probabilistically determined based on the alternative propensity as well as the propensity of all other alternatives. Thus, exogenous variables directly affect all alternatives. The approach is well suited to accommodate for large number of alternatives without a sizable increase in computational burden. The model was estimated using crash data at Traffic Analysis Zone (TAZ) level from Florida. The modeling results clearly illustrate the applicability of the proposed framework for crash proportion analysis. Further, the Excess Predicted Proportion (EPP)—a screening performance measure analogous to Highway Safety Manual (HSM), Excess Predicted Average Crash Frequency is proposed for hot zone identification. Using EPP, a statewide screening exercise by the various vehicle types considered in our analysis was undertaken. The screening results revealed that the spatial pattern of hot zones is substantially different across the various vehicle types considered.

1. Introduction

The Federal, State, and Local government officials and transportation engineers have been working with consistent efforts in reducing both road crash related fatalities and fatality rates. In the United States, traffic collisions have steadily declined from 2003 to 2011. However, traffic fatalities and fatality rates rose in 2012 with slight drop in 2013 highlighting the challenges faced by the safety community. Overall 29,989 people were reported to be killed in the United States from traffic crashes in 2014 (NHTSA, 2016). These facts highlight that there is a need for continued efforts to identify remedial measures to reduce crash occurrences and crash consequences. Towards this end, the traffic safety literature has evolved along two main dimensions: collision frequency analysis and collision severity analysis. The former group of studies is focused on identifying factors that result in traffic collisions while the latter group is concentrated on ameliorating the consequences of traffic crashes (conditional on their occurrences). The current study contributes to traffic safety literature along the first dimension—identify factors that inform us about traffic collision occurrences.

Collision frequency analysis is traditionally undertaken at the microscopic and macroscopic levels. The microscopic safety analysis focuses on roadway entities such as segments, intersections, and corridors (Lee et al., 2017). The studies broadly aim to identify contributing factors for traffic crashes from roadway geometric design, traffic characteristics, and provide specific engineering countermeasures to alleviate traffic collisions. On the other hand, the macroscopic safety analysis relates traffic crashes aggregated at a spatial level (traffic analysis zone (TAZ), census tract or county) with demographic, socioeconomic, built environment, traffic attributes and/or roadway characteristics at a study unit level. While microscopic level analysis is more focused on the engineering design and evaluation, the macroscopic analysis provides a broad spectrum for long-term policy based countermeasures such as enactments of traffic laws, police enforcement, education, and area-wide engineering solutions (Lee et al., 2014b). There has been growing recognition within the planning community to incorporate macroscopic models as part of long range transportation plans. For example, the Moving Ahead for Progress in the 21st Century Act (MAP-21) and Fixing America's Surface Transportation Act (FAST) have emphasized the role of macro-level safety analysis in planning.

* Corresponding author.
E-mail address: jaeyoung@knights.ucf.edu (J. Lee).

https://doi.org/10.1016/j.aap.2017.11.017
Received 6 August 2015; Received in revised form 21 September 2017; Accepted 13 November 2017
Available online 20 November 2017

While total traffic fatalities and fatality rates show a downward trend, NHTSA statistics indicate that the proportions of motorcycles, bicycles, and pedestrians in fatal crashes have considerably increased whereas the proportion of passenger cars has decreased from 2005 to 2014 (NHTSA, 2016). The proportions of motorcyclists and non-motorists involved in fatal crashes has risen from 11% to 14% and 13% to 18%, respectively, since the last decade. Despite these increases, earlier studies in the safety area have predominantly focused either on total crashes or crashes involving passenger cars/trucks or crashes involving non-motorists. These studies provide important information in improving safety situation for different road user groups separately. However, it is also important to examine critical factors contributing to crash occurrences including all road user groups in a single framework, which would allow stakeholders to devise a more general safety conscious planning. Towards that end, the main objective of this study is to explore the proportions of traffic crashes at TAZ level across different vehicle types including both motorized and non-motorized group of road users. Specifically, the current study considers the proportion of crashes by vehicle type as the dependent variable and estimates a TAZ level mixed multinomial fractional split model. The proposed approach will assist transportation planners and engineers in devising safety conscious plans. Specifically, traffic engineers and planners can understand the factors affecting the proportions of crashes by vehicle types from the modeling results, and use these findings to design long-term transportation plans. Furthermore, the predicted proportions of crashes from the proposed model can be used to identify hotspots for each vehicle type. Thus, traffic engineers and planners can proactively provide effective safety countermeasures for the zones with excessive high proportions for specific vehicle types. The vehicle types considered in our analysis include: passenger car,[1] van, light truck, medium and heavy truck, bus, motorcycle, bicycle, and pedestrian. The reader would note that the model employed is not similar to the traditional multinomial logit model because the dependent variable in our case is proportion by vehicle type whereas it is a single chosen alternative in the multinomial logit model.

1.1. Literature review

The macro-level safety studies have been conducted by total crashes, crashes by severity levels (such as no injury, minor injury, severe injury, and fatal injury) and crashes by vehicle type (such as motor vehicle, pedestrian, and bicycle). It is beyond the scope of this paper to exhaustively review all the studies in frequency modeling (see Lord and Mannering (2010) and Yasmin and Eluru (2016) for a detailed review). In our study, we group literature in the context of our research effort along two main groups: 1) independent frequency models for a single dependent variable or multiple dependent variables; and 2) multivariate count models for the multiple dependent variables are estimated.

In the first group of studies, usually either total number of crashes in the study unit or crashes by vehicle type or severity level are investigated. In some studies, multiple dependent variables are considered in the analysis while ignoring the relationship across the dependent variables. For example, Noland and Quddus (2004) developed fixed-effect negative binomial models for bicycle and pedestrian crashes for severe and minor injury. Nevertheless, these models did not account for the possible correlation among dependent variables. Lee et al. (2013) estimated a series of negative binomial models for total, severe, driving under the influence (DUI), pedestrian and bicycle crashes based on zero-inflated Poisson (ZIP) framework while ignoring the relationships across the dependent variables. Abdel-Aty et al. (2013) analyzed the contributing factors for total, and severe pedestrian crashes, using negative binomial models without considering common unobserved factors across different crash types.

In the second group of studies, multivariate models that recognize the dependencies between multiple dependent variables are estimated. For example, Song et al. (2006) analyzed intersection, intersection-related, driveway access, and non-intersection crashes at the county level. The authors developed Bayesian multivariate conditional autoregressive models that can account for the dependencies between the variables and spatial effects. (Narayanamoorthy et al., 2013) explored pedestrian and bicycle crash frequencies by injury severity levels. The authors adopted a multivariate model to accommodate jointness in the dependent variables, while considering spatial dependence effects. Furthermore, Lee et al. (2015b) estimate multivariate Poisson lognormal models to analyze motor-vehicle, pedestrian, and bicycle crashes. The authors found that the multivariate model accounting for unobserved common factors across the dependent variables outperforms the univariate model that does not take the dependencies between dependent variables into consideration. Recently, Nashad et al. (2016) adopted a bivariate copula based modeling approach to examine pedestrian and bicyclist crashes simultaneously.

Based on earlier literature, the factors that are likely to affect motor vehicle crashes are socio-demographic factors, land-use characteristics, roadway-related variables, and traffic characteristics (Kim et al., 2006; Lee et al., 2015b). Based on 0.1 mi^2 grid structure, Kim et al. (2006) estimated various regression models based on the grid. The authors identified that population, total job, total economic output, and commercial area had positive relationship with motor vehicle crashes. Lee et al. (2015b) analyzed motor vehicle crashes based on TAZs. The authors found that population, total employment, total economic output, commercial area, roadway types, proportion of households without a vehicle, number of accommodation facilities per square mile, and number of traffic signals per mile have a positive effect on motor vehicle crashes. Very few researchers have explored truck-involved crashes at a macro-level. Pasupuleti and Pulugurtha (2013) identified several zonal characteristics related to truck crashes. The authors revealed that truck crashes are positively correlated to industrial areas and areas with large residential lots but negatively correlated with highly populated areas.

Several factors that tend to increase bicycle crashes have been identified by researchers. Noland and Quddus (2004) found that national health service staff per population, percentage of motorway, percentage of trunk road density, percentage of older vehicle, percentage of households without cars, per capita expenditure on alcohol, population, percentage of population aged 65 or over have a positive relationship with severe bicycle crashes. On the other hand, length of inpatient stay in the hospital, income level, percentage of population aged 45–64 have a negative relationship with severe bicycle crashes. Kim et al. (2006) attempted several candidate variables including population, total job, total economic output, hospital, park, commercial area, and school for bicycle crashes but only population was found significant and positively related to bicycle crashes. Lee et al. (2013) investigated the residence of bicyclists who were involved in traffic crashes. It was revealed that median age, average travel time to work, household income, and workers in the primary industry field were negatively associated with the number of crash involved bicyclists. On the other hand, Hispanic people, workers commuting by bicycle, urban area, and older buildings were positively associated with the number of bicyclists who were involved in traffic crashes. Lee et al. (2015b) developed a multivariate model for motor vehicle-to-vehicle, bicycle-to-vehicle, and pedestrian-to-vehicle crashes. The authors uncovered that vehicle-miles-traveled, population, commuters using bicycle, hotel/motel/timeshare rooms per square mile, employments and school enrollments per square mile, number of traffic signals per mile are likely to increase bicycle crashes whereas proportion of roadway with speed limit of 20 mph or less tends to decrease bicycle crashes. The above mentioned literature suggest that the factors that tend to increase bicycle crashes are population, distance to urban location, employment,

[1] Passenger cars include sedans and sport utility vehicles (SUV).