JOURNAL OF TRAFFIC AND TRANSPORTATION ENGINEERING (ENGLISH EDITION) 2016; 3 (3): 236–242

Available online at www.sciencedirect.com

ScienceDirect

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

Original Research Paper

Random-parameters analysis of highway characteristics on crash frequency and injury severity

Bismark R. D. K. Agbelie*

School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 28 April 2015
Received in revised form
25 September 2015
Accepted 29 September 2015
Available online 29 April 2016

Keywords:
Random-parameters
Multinomial logit model
Negative binomial
Heterogeneity
Highway crash
Highway segment
Vertical grade
Stop sign
Marginal effects

ABSTRACT

Factors that affect highway-related crash frequency and injury severity vary across observations. Using a methodology that does not account nor correct for heterogeneity in observed and unobserved crash factors across highway segments may lead to biased and inconsistent estimated coefficients, thus resulting in erroneous inferences. The present paper demonstrates the use of random-parameters models to facilitate and enhance how crash factors affect crash frequency and injury severity along a highway segment. The results indicate that a unit increase in the presence of stop sign along a highway segment reduces crash frequency by 2.471 for 87.24% of the roadway segments. For the remaining 12.76% of the roadway segments, crash frequency is increased by the same margin. Using the random-parameters multinomial logit model, the result indicates that, for 90.89% of the observations, the presence of a stop sign on a highway segment increases the probability of the injury outcome. For 9.11% of the observations, the presence of a stop sign on a highway segment reduces the probability of the injury outcome, and the marginal effect value across observations is 0.0017. Vertical grades greater than 5% increase crash frequency for 58.46% of the highway segments, and decrease for 41.54% of the highway segments by 0.121 for one unit increase in vertical grades.

1. Introduction

In past decades, highway traffic safety has seen some improvements due to enhanced vehicle safety features and highway design. Nonetheless, crash frequencies in many countries have not significantly decreased as expected. Recent worldwide highway safety reports indicate that, out of every single road traffic fatality, twenty people would be injured, and by 2030, highway-related crashes will be among the top five leading causes of death (WHO, 2013). Worldwide highway crash statistics worldwide show that about 1.24 billion people are killed per year, with over 20–50 billion people injured, costing over $500 billion in total (WHO, 2013). Fatalities and

* Corresponding author. Tel.: +1 765 409 8800.
E-mail address: bagbelie@purdue.edu.

Peer review under responsibility of Periodical Offices of Chang'an University.
http://dx.doi.org/10.1016/j.jtte.2015.09.006

injuries have a devastating effect on families, often resulting in distress, productivity loss, and huge medical costs.

While numerous research efforts have been conducted to investigate and comprehend the factors influencing crash frequency on highway segments and intersections to provide effective countermeasures, much work is still needed (Zohar, 2010). The purpose of this paper is to demonstrate an approach that improves the understanding of the influence of identified factors in predicting crash frequency and injury severity along highway segments.

A number of methodological approaches have been employed in recent years to understand the influencing factors of highway traffic safety. Out of these methodological approaches, the Poisson and negative binomial models have been commonly adopted to predict crash frequencies, and the multinomial logit model has been used to predict injury severity, with respect to modeling count and discrete data in traffic highway safety (Agbelie, 2016; Agbelie and Roshandeh, 2014; Simon and Corbett, 1996; Simons-Morton and Hartos, 2003). Additionally, to account for the presence of "no crash" and other variations in the dataset, some other modified count models were developed, including the Poisson and negative binomial zero-inflated models, Conway–Maxwell–Poisson generalized linear models, and Poisson and negative binomial with random effects models (Lord et al., 2008; Shankar et al., 2003).

Even though a number of studies (Cameron and Trivedi, 1986; Eluru et al., 2008; Fridstrøm et al., 1995) adopted count-data and discrete models to enhance the understanding of many factors influencing highway-related safety, the underlining assumption of the most models was that the estimated parameters fixed across observations for intersections or highway segments. If the estimated parameters do actually vary across observations, and the methodology fixes these parameters, it will lead to biased and inconsistent coefficient estimates (Greene, 2012; Washington et al., 2011).

Accounting for the possibility of unobserved heterogeneity across observations, by allowing all or some of the estimated parameters to vary across intersections or highway segments, can improve the current comprehension of the influencing factors that impact highway-related safety. The application of a random-parameters model to account for unobserved heterogeneity across observations has been established to have enormous potential, as evidenced by previous studies on discrete outcome models (McFadden and Train, 2000; Milton et al., 2008). Thus, given the possibility of heterogeneity in observed and unobserved factors for highway segment crash frequency and severity data, the random-parameters approach can be a suitable methodology to adopt.

2. Empirical setting

The random-parameters negative binomial model is demonstrated using detailed data from Washington State. The crash, environmental, highway inventory, and traffic volume data for Washington State were made available from the Highway Safety Information System. The data consist of 21,069 highway segments in the year 2011, with each highway segment presenting homogenous features to identify the beginning and end points. The minimum and maximum lengths of the segments are 0.12, and 4.8 miles respectively, with a standard deviation of 0.19 miles. The crash data details available to the present study include the total number of crashes on a road segment, crash type (e.g., 27 detailed descriptions including head-on collisions), and severity of crash (e.g., property damage, injury and fatality). The highway inventory dataset available includes the functional class of road, segment length, gradient type (e.g., vertical curve, angle point or gap), direction of grade, percent of grade, access control, shoulder width and type, median width and type, terrain (e.g., rolling, mountainous, and level), lane type and number, pavement surface condition, and lighted condition. The traffic flow characteristics include the average annual daily passenger traffic volume, percentage of combination truck traffic, average annual daily combination truck traffic volume, intersection control type, and posted speed limit. The other remaining group of variables includes the weather and year/month of crash. As a result of the quantum of data made available, only the summary statistics of the variables statistically significant in the subsequent estimation of the random-parameters negative binomial model are discussed. The descriptive statistics of the selected continuous variables are presented in Table 1, while the distributions of the indicator variables are presented in Table 2. For example, in Table 1, the mean, standard deviation, minimum, and maximum values for average annual daily combination truck traffic volume are 5002, 5083, 205, and 20,115, respectively. In Table 2, highways with a vertical grade lower than 5% constituted 73% of the model data, while highways with a vertical grade greater than 5% constituted 27% of the model data. The crash injury severity distributions are as

Table 1 – Descriptive statistics of selected continuous variables.

Variable description	Mean	Std. dev.	Minimum	Maximum
Number of crashes	4.964	3.881	1	20
Average annual daily combination truck traffic volume	5002	5083	205	20,115
Lane width (in feet)	10.328	1.190	9	12
Median width (in feet)	23.418	66.306	0	99
Average annual daily passenger car traffic volume	61,336	57,634	1568	209,827

Table 2 – Descriptive statistics of selected indicator variables.

Variable description (1 if yes, 0 otherwise)	Distribution (%)	
	0	1
Indicator variable: road segment in urban area	95	5
Indicator variable: rolling terrain segment	98	2
Presence of ice on pavement surface	97	3
Crash occurring during daylight	92	8
Presence of stop signs	94	6
Vertical grade greater than 5%	73	27