Accident Analysis and Prevention 98 (2017) 223–231

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

The effect of zonal factors in estimating crash risks by transportation modes: Motor vehicle, bicycle and pedestrian

Jie Wang[a], Helai Huang[a,*], Qiang Zeng[b]

[a] School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan, China
[b] School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, Guangdong, China

ARTICLE INFO

Article history:
Received 29 June 2016
Received in revised form 9 October 2016
Accepted 13 October 2016
Available online 19 October 2016

Keywords:
Transportation modes
Macroscopic variables
Unobserved heterogeneity
Buffer width
Intersection safety

ABSTRACT

Objectives: This paper aimed to (i) differentiate the effects of contributory factors on crash risks related to different transportation modes, i.e., motor vehicle, bicycle and pedestrian; (ii) explore the potential contribution of zone-level factors which are traditionally excluded or omitted, so as to track the source of heterogeneous effects of certain risk factors in crash-frequency models by different modes.
Methods: Two analytical methods, i.e. negative binomial models (NB) and random parameters negative binomial models (RPNB), were employed to relate crash frequencies of different transportation modes to a variety of risk factors at intersections. Five years of crash data, traffic volume, geometric design as well as macroscopic variables at traffic analysis zone (TAZ) level for 279 intersections were used for analysis as a case study.
Results: Among the findings are: (1) the sets of significant variables in crash-frequency analysis differed for different transportation modes; (2) omission of macroscopic variables would result in biased parameters estimation and incorrect inferences; (3) the zonal factors (macroscopic factors) considered played a more important role in elevating the model performance for non-motorized than motor-vehicle crashes; (4) a relatively smaller buffer width to extract macroscopic factors surrounding the intersection yielded better estimations.

1. Introduction

Many communities have increased their interest in the implementation of multimodal transportation and advocated for the shift from motor vehicles to non-motorized modes of transportation, i.e., walking and cycling. In spite of the health and environmental benefits, an increasing number of crashes involving pedestrians and bicyclists has become a major concern in improving traffic safety. For example in 2013, the United States had 4735 pedestrian and 743 bicyclist deaths, accounting for 18% of all U.S. highway fatalities (NHTSA, 2013). The Federal Highway Administration's office of safety has established pedestrian and bicyclist safety as one of its top priorities. Thus, it is essential for traffic safety engineers to provide appropriate countermeasures or policies to achieve friendly and safe multimodal transportation.

A comprehensive understanding of contributing factors associated with crash occurrences by different modes is a prerequisite for developing safety improvement programs to effectively reduce traffic crashes. For a given road entity (e.g. road segments or intersections), the potential factors associated with multimodal crashes could be summarized as in Fig. 1, according with Miranda-Moreno et al. (2011), Mitra and Washington (2012), Ukkusuri et al. (2012) and Strauss et al. (2013, 2014). The factors influencing road-entity-level crash frequency by modes include macroscopic factors related to built environment of the road entities – such as population and economic characteristics, land use characteristics and travel behaviors – as well as road features and traffic characteristics of the road entities. In addition, crash occurrence is also associated with individual characteristics such as gender, age, education, alcohol consumption, and other driver and pedestrian behaviors (Ryb et al., 2007). Although discrete individual-level factors are not available to be integrated into the crash-frequency model, individual characteristics are always influenced by macroscopic factors (Christoffel and Gallagher, 1999). Therefore, macroscopic factors could serve as a surrogate for individual behaviors.

The choice of appropriate analytical method and the selection of representative explanatory variables are two important considerations for obtaining accurate model predictions. Over the past three decades, considerable research efforts have been devoted to

* Corresponding author.
E-mail addresses: jie_wang@csu.edu.cn (J. Wang), huanghelai@csu.edu.cn (H. Huang), zengqiang@scut.edu.cn (Q. Zeng).

http://dx.doi.org/10.1016/j.aap.2016.10.018

developing and applying sophisticated methodological approaches associated with the analysis of crash frequency. Detailed descriptions and assessments of crash-frequency models can be found in the review papers by Lord and Mannering (2010) and Mannering and Bhat (2014). However, relatively few studies have focused on the identification and inclusion of traditionally excluded or omitted variables in crash-frequency analysis. In particular, variables related to macroscopic factors previously described (in Fig. 1) are normally unavailable in crash databases and as a result have rarely been examined in great detail. Mitra and Washington (2012) is one of a few studies exploring the omitted variables in crash-frequency modeling. The authors developed two different models of estimating intersection crash frequency, one with traffic volume as the only independent variable, and the other with several spatial factors in addition to commonly included geometric design and traffic factors. Through contrastive analysis of the two models, results indicated that some spatial factors, such as local influences of weather, sun glare, proximity to drinking establishment, proximity to school and demographic attributes near intersections, have significant explanatory power and their exclusion leads to biased estimates.

Statistical methods such as spatial and temporal correlation, multilevel, random effect, random parameter, and latent class approaches have been developed to address this issue of unobserved heterogeneity (Anastasopoulos and Mannering, 2009; Dong et al., 2016; Mannering et al., 2016; Quddus, 2008; Wang and Huang, 2016; Xu and Huang, 2015; Xu et al., 2016), as these omitted explanatory variables can be regarded as part of the unobserved heterogeneity. Unobserved heterogeneity impacts traffic safety analysis in two ways: the first problem is that the selected explanatory variables cannot fully account for the cross-section or longitudinal-section variations in crash counts due to unobserved road geometrics, environmental factors, driver behavior and other confounding factors, which lead to impaired predictive performance of the model (called heterogeneity in model prediction); the second problem is that these unobserved factors are always correlated with observed factors and thus biased parameters will be estimated and incorrect inferences could be drawn (called heterogeneity in the coefficient estimator). While these approaches will mitigate the adverse impacts of omitting significant explanatory variables, the resulting model estimates still fail to track the original source of heterogeneity and quantify the safety effect of omitted variables (such as macroscopic factors shown in Fig. 1). Omission of important explanatory variables still remains a problem even with advanced statistical approaches to capture unobserved heterogeneity (Mannering et al., 2016).

The study by Mitra and Washington (2012) attempted to investigate the safety effect of some important omitted variables on total crash frequency and their contribution on model estimation. As Venkataraman et al. (2013) stated, frequency models of crash outcome type can provide substantial insights into the effect of explanatory variables and assist in examining the heterogeneity effects in roadway geometric features. This paper aims to extend previous research (Mitra and Washington, 2012) and investigate how macroscopic factors affect the crash-frequency analysis for different transportation modes. This is because there may be some inconsistent impacts of some macroscopic variables on motor vehicle and non-motorized (including bicycle and pedestrian) crashes. For example, Lee et al. (2015) utilized a multivariate model for investigating motor vehicle and non-motorized crashes at the macroscopic level. Results for the parameter estimation suggested that some zonal variables related to demographics and road characteristics have different directional effects on motor vehicle and non-motorized crashes. Meanwhile, the most appropriate width of buffer to extract macroscopic factors may be inconsistent between modeling motor vehicle and non-motorized crashes. Therefore, it is advisable to model the crash frequency by separate transportation modes to examine the effects of macroscopic factors.

In summary the objective of this paper is twofold: (1) to examine the effects of a host of contributing factors including both macroscopic and microscopic factors on crash occurrence with respect to different transportation modes; (2) to shed further light on the contribution of the macroscopic factors which are traditionally excluded or omitted variables, to tracking the source of heterogeneity effects in coefficient estimators of regularly used variables and improving the model performance in crash-frequency analysis related to different modes.

2. Data preparation

In this study, data collected for 279 intersections located in Hillsborough County, Florida, USA were used to develop the intersection crash-frequency models for different transportation modes. The data for the analysis was mainly divided into four types: traffic crash data, traffic characteristics, road characteristics related to geometric design, traffic control/regulatory of the intersection, macroscopic factors including trip production/attraction, demographic and socio-economic characteristics surrounding the intersection. The derivation and processing of these data sources are described next.

2.1. Crash data

Crash data for the intersections in a five-year period (2005–2009) were obtained from the Florida Department of Transportation (FDOT) Crash Analysis Reporting (CAR) system. Crashes were categorized as intersection related crashes if they occurred within the curb-line limits of the intersection or if they occurred within the influence area of the intersection, which is 250 feet away from the stop line. Intersection-level crash data was disaggregated into motor vehicle, bicycle, and pedestrian crashes. A motor vehicle crash was defined as a collision between two or more motor vehicles or between a motor vehicle and an object. A bicycle crash referred to a collision between a motor vehicle and a bicycle. Likewise, a pedestrian crash denoted a collision between a motor vehicle and a pedestrian.

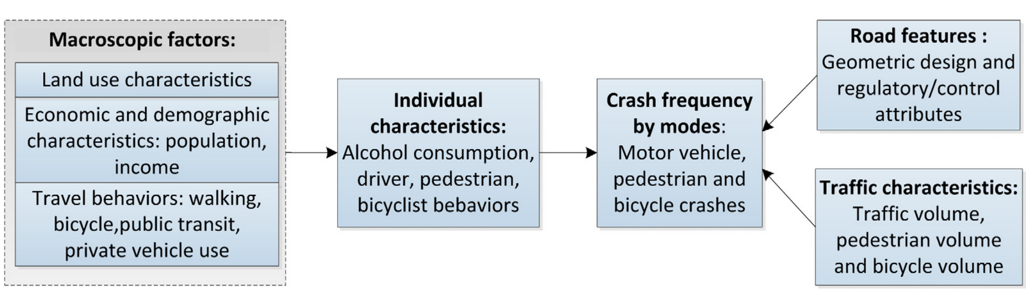

Fig. 1. Factors associated with multimodal crashes.