



Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models



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

Department of Civil, Environment and Construction Engineering, University of Central Florida, Orlando, FL 32816, USA

ARTICLE INFO

Article history:

Received 6 August 2015

Received in revised form 9 December 2015

Accepted 15 April 2016

Available online 3 May 2016

Keywords:

Macro-level crash analysis

Pedestrian and bicycle crashes

Dual-state models

Spatial independent variables

Zero-inflated negative binomial

Hurdle negative binomial models

ABSTRACT

This study attempts to explore the viability of dual-state models (i.e., zero-inflated and hurdle models) for traffic analysis zones (TAZs) based pedestrian and bicycle crash frequency analysis. Additionally, spatial spillover effects are explored in the models by employing exogenous variables from neighboring zones. The dual-state models such as zero-inflated negative binomial and hurdle negative binomial models (with and without spatial effects) are compared with the conventional single-state model (i.e., negative binomial). The model comparison for pedestrian and bicycle crashes revealed that the models that considered observed spatial effects perform better than the models that did not consider the observed spatial effects. Across the models with spatial spillover effects, the dual-state models especially zero-inflated negative binomial model offered better performance compared to single-state models. Moreover, the model results clearly highlighted the importance of various traffic, roadway, and sociodemographic characteristics of the TAZ as well as neighboring TAZs on pedestrian and bicycle crash frequency.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Active forms of transportation such as walking and bicycling have the lowest impact on the environment and improve the physical health of pedestrians and bicyclists. With growing concern of worsening global climate change and increasing obesity among adults in developed countries, it is hardly surprising that transportation decision makers are proactively encouraging the adoption of active forms of transportation for short distance trips. However, transportation safety concerns related to active transportation users form one of the biggest impediments to their adoption as a preferred alternative to private vehicle use for shorter trips. According to the National Highway Traffic Safety Administration (NHTSA), from 2004 to 2013, the proportion of pedestrian fatalities has steadily increased from 11% to 14% (NHSTA, 2013a) while the proportion of bicyclist fatalities has increased from 1.7% to 2.3% (NHTSA, 2013b). Thus, traffic crashes and the consequent injury and fatality remain a deterrent for active modes of transportation, specifically in North American communities (Wei and Lovegrove, 2013). Any effort to reduce the social burden of these crashes would necessitate the implementation of policies that enhance safety for active transportation users. An important tool to

identify the critical factors affecting occurrence of bicycle crashes is the application of planning level crash prediction models.

Traditionally, transportation crash prediction models are developed for two levels: micro and macro-level. At the micro-level, crashes on a segment or intersection are analyzed to identify the influence of geometric design, lighting and traffic flow characteristics with the objective of offering engineering solutions (such as installing sidewalk and bike lane, adding lighting). On the other hand, the macro-level crashes from a spatial aggregation (such as traffic analysis zone (TAZ) or county) are considered to quantify the impact of socioeconomic and demographic characteristics, transportation demand and network attributes so as to provide countermeasures from a planning perspective. The current research effort contributes to burgeoning literature on active transportation user safety by examining pedestrian and bicycle crashes in the state of Florida at a macro-level. Specifically, in this study, a comprehensive analysis of pedestrian and bicycle crashes is conducted at the macro-level by employing several crash frequency models. A host of exogenous variables including socio-economic and demographic characteristics, transportation network characteristics, and traffic flow characteristics are considered in the model development. In addition, exogenous variables from neighboring zones are also considered in the analysis to account for spatial proximity effects on crash frequency. The overall model development exercise will allow us to identify important determinants of pedestrian and bicycle crashes in Florida while also providing valuable

* Corresponding Author.

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

insight on appropriate model frameworks for macro-level crash analysis.

1.1. Literature review

A number of research efforts have examined transportation (vehicle, pedestrian and bicycle) related crash frequency (see (Lord and Mannering, 2010) for a detailed review). These studies have been conducted for different modes (vehicle, pedestrian and bicycle) and for different scales – micro (such as intersection and segment) and macro-level (such as census tract, traffic analysis zone, county). The model structures considered in earlier literature include Poisson, Poisson–Lognormal, Poisson–Gamma regression (also known as negative binomial (NB)), Poisson–Weibull, and Generalized Waring models (Abdel-Aty and Radwan, 2000; Miaou et al., 2003; Aguero-Valverde and Jovanis, 2008; Lord and Miranda-Moreno, 2008; Maher and Mountain, 2009; Cheng et al., 2013; Peng et al., 2014). Among these model structures, the NB model offers a closed form expression while relaxing the equal mean variance equality constraint and serves as the workhorse for crash count modeling.

1.2. Handling excess zeros

One methodological challenge often faced in analyzing count variables is the presence of a large number of zeros. The classical count models (such as Poisson and NB) allocate a probability to observe zero counts, which is often insufficient to account for the preponderance of zeros in a count data distribution. In crash count variable models, the presence of excess zeros may result from two underlying processes or states of crash frequency likelihoods: crash-free state (or zero crash state) and crash state (see (Shankar et al., 1997) for more explanation). The zero crash state can be a mixture of true zeros (where the zones are inherently safe (Shankar et al., 1997)) and sampling zeros (where excess zeros are results of potential underreporting of crash data (Miaou, 1994)). In presence of such dual-state, application of single-state model (Poisson and NB) may result in biased and inconsistent parameter estimates.

In econometric literature, two potential relaxations of the single-state count models are proposed for addressing the issue of excess zeros. The first approach – the zero inflated (ZI) model – is typically used for accommodating the effect of both true and sampling zeros, and has been employed in several transportation safety studies (Shankar et al., 1997; Chin and Quddus, 2003). The second approach – the Hurdle model – is typically used in the presence of sampling zeros and has seldom been used in transportation safety literature. The two approaches differ in the approach employed to address the excess zeros. The appropriate framework for analysis might depend on the actual empirical dataset under consideration. Table 1 presents a summary of previous studies that have considered zero-inflated and hurdle models to analyze crashes. The table provides information on type and severity of crash analyzed, spatial and temporal unit of analysis and the data collection duration. From the table, it is evident that all the existing zero-inflated and hurdle studies are conducted at a micro-level such as segment and intersection except for Brijs et al. (2006), which conducted crash analysis at macro-level by assigning crashes to the closest weather station. Second, with the exception of study (Hu et al., 2011; Hosseinpour et al., 2013; Hosseinpour et al., 2014), the range of observation of the study period is one year or less; that may explain the preponderance of zeros in the data (Lord et al., 2005). Third, the zero-inflated model always offers better statistical fit to crash data.

1.3. Issues with dual-state models

To be sure, several research studies have criticized the application of zero-inflated model for traffic safety analysis (Lord et al.,

2005; Lord et al., 2007; Kweon, 2011). The authors question the basic dual-state assumption for crash occurrence and have conducted extensive analysis at the micro-level and indicated that the development of models with dual-state process is inconsistent with crash data at the micro-level. While the reasoning behind the “non-applicability” is plausible for micro-level the reasoning does not necessarily carry over to the macro-level crash counts. At the macro-level it is possible to visualize dual-state data generation with some macro-level units having zero pedestrian and bicyclist crashes – possibly because these spatial units have no pedestrian and bicycle demand (because of lack of walking and cycling infrastructure). In such cases the dual-state representation will allow us to identify spatial units that are likely to have zero cases as a function of exogenous variables (for example very low walking and cycling infrastructure might result in the higher probability of a zero state). Hence, we have considered the possible existence of dual-state models for pedestrian and bicycle crashes at the macro level in our research. If the data generation does support the dual-state models, ignoring the excess zeros and estimating traditional NB models will result in biased estimates.

1.4. Spatial spillover effects

In macro-level analysis, crashes occurring in a spatial unit are aggregated to obtain the crash frequency. The aggregation process might introduce errors in identifying the exogenous variables for the spatial unit. For example, a crash occurring closer to the boundary of the unit might be strongly related to the neighboring zone than the actual zone where the crash occurred. This is a result of arbitrarily demarcating space. To accommodate for such spatial unit induced bias, two approaches to incorporate spatial dependency are considered: (1) spatial error correlation effects (unobserved exogenous variables at one location affect dependent variable at the targeted and neighboring locations) and (2) spatial spillover effects (observed exogenous variables at one location having impacts on the dependent variable at both the targeted and neighboring locations) (Narayanamoorthy et al., 2013). Several research efforts have accommodated for spatial random error in safety literature (for example see (Huang et al., 2010; Siddiqui et al., 2012; Lee et al., 2015)). On the other hand, researchers have considered a spatially lagged dependent variable at neighboring units for the spatial spillover effects (Lascaia et al., 2000; Quddus, 2008; Ha and Thill, 2011). However, the utility of such spatially lagged dependent variable models, particularly for prediction, is limited since developing prediction frameworks for spatially lagged models is involved. In our analysis, to accommodate for spatial effects, we propose the consideration of exogenous variables from neighboring zones for accounting for spatial dependency. The approach, referred to as spatial spillover model, is easy to implement and allows practitioners to understand and quantify the influence of neighboring units on crash frequency.

In summary, the current study contributes to non-motorized macro-level crash analysis along two directions: (1) evaluate the viability of dual-state models for non-motorized crash analysis at macro-level; and (2) introduction of spatial independent variables accounting for spatial spillover effects on crash frequency. Towards this end, conventional single-state model (i.e., NB) and two dual-state models (i.e., zero-inflated NB (ZINB) and hurdle NB (HNB)) with and without spatial independent variables are developed for both pedestrian and bicycle crashes at a TAZ level in Florida. Overall, both pedestrian and bicycle crashes have 6 model structures estimated – NB model without/with spatial effects (aspatial/spatial NB), ZINB model without/with spatial effects (aspatial/spatial ZINB), and HNB model without/with spatial effects (aspatial/spatial HNB). The model development process considers a sample for model calibration and a hold-out sample

Download English Version:

<https://daneshyari.com/en/article/571941>

Download Persian Version:

<https://daneshyari.com/article/571941>

[Daneshyari.com](https://daneshyari.com)