Journal of Safety Research 46 (2013) 67–76

Contents lists available at SciVerse ScienceDirect

Journal of Safety Research

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

Effect of driver's age and side of impact on crash severity along urban freeways: A mixed logit approach

Kirolos Haleem *, Albert Gan [1]

Department of Civil and Environmental Engineering, Florida International University, 10555 West Flagler Street, EC 3680, Miami, FL 33174, United States

ARTICLE INFO

Article history:
Received 7 July 2012
Received in revised form 13 February 2013
Accepted 18 April 2013
Available online 1 May 2013

Keywords:
Mixed Logit
Severity
Urban Freeway
Driver's Age
Side of Impact

ABSTRACT

Introduction: This study identifies geometric, traffic, environmental, vehicle-related, and driver-related predictors of crash injury severity on urban freeways. *Method:* The study takes advantage of the mixed logit model's ability to account for unobserved effects that are difficult to quantify and may affect the model estimation, such as the driver's reaction at the time of crash. Crashes of 5 years occurring on 89 urban freeway segments throughout the state of Florida in the United States were used. Examples of severity predictors explored include traffic volume, distance of the crash to the nearest ramp, and detailed driver's age, vehicle types, and sides of impact. To show how the parameter estimates could vary, a binary logit model was compared with the mixed logit model. *Results:* It was found that the at-fault driver's age, traffic volume, distance of the crash to the nearest ramp, vehicle type, side of impact, and percentage of trucks significantly influence severity on urban freeways. Additionally, young at-fault drivers were associated with a significant severity risk increase relative to other age groups. It was also observed that some variables in the binary logit model yielded illogic estimates due to ignoring the random variation of the estimation. Since the at-fault driver's age and side of impact were significant random parameters in the mixed logit model, an in-depth investigation was performed. It was noticed that back, left, and right impacts had the highest risk among middle-aged drivers, followed by young drivers, very young drivers, and finally, old and very old drivers. *Impact on Industry:* To reduce side impacts due to lane changing, two primary strategies can be recommended. The first strategy is to conduct campaigns to convey the hazardous effect of changing lanes at higher speeds. The second is to devise in-vehicle side crash avoidance systems to alert drivers of a potential crash risk. *Conclusions:* The study provided a promising approach to screening the predictors before fitting the mixed logit model using the random forest technique. Furthermore, potential countermeasures were proposed to reduce the severity of impacts.

1. Introduction

High-speed roadways such as freeways have continued to be a research focus due to their high correlation with injury severity (Malyshkina & Mannering, 2008; Renski, Khattak, & Council, 1999). Identifying the significant predictors of crash injury severity along these facilities can help to select more effective countermeasures that can better tackle the underlying safety deficiencies. This paper proposes to apply the mixed logit model as a more robust approach to modeling injury severity. The model has been shown to be effective in analyzing injury severity (Anastasopoulos & Mannering, 2011; Gkritza & Mannering, 2008; McFadden & Train, 2000; Milton, Shankar, & Mannering, 2008; Pai, Hwang, & Saleh, 2009). The model is characterized by its capability to account for unobserved predictors, which is a highly desirable property in severity studies due to the difficulty in quantifying some features, such as the driver behavior and reaction at the time of a crash (see Kim, Ulfarsson, Kim, & Shankar, 2011).

As demonstrated by McFadden and Train (2000) and Train (2009), the mixed logit model, also referred to as the mixed multinomial logit model, the random parameters logit model, and the error-components logit model, is a highly flexible model that can be used to approximate different random utility functions. The model can account for the standard multinomial logit model limitations, while allowing for random variation across the observations and unrestricted substitution patterns, as well as correlation in unobserved features across time. In this model, some parameters are held fixed while others are allowed to be random. The random effects mixed logit model is preferred to the fixed effects model since it allows time-invariant explanatory variables, such as gender and age, to be used as main effects.

The application of the mixed logit model in analyzing traffic crash injury severity on urban freeways that typically carry a high number of commuters has not been extensively explored in safety studies. According to the Traffic Safety Facts (National Highway Traffic Safety Administration, 2007), urban fatalities increased by 8% from 1998 to

* Corresponding author. Tel.: +1 321 276 7889; fax: +1 305 348 2802.
E-mail addresses: khaleemm@fiu.edu (K. Haleem), gana@fiu.edu (A. Gan).
[1] Tel.: +1 305 348 3116; fax: +1 305 348 2802.

0022-4375/$ – see front matter © 2013 National Safety Council and Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jsr.2013.04.002

2007, while rural fatalities decreased by 9% during the same time span; the significance of this increase is highlighted by the fact that, nationwide, Florida was ranked second after California for urban fatalities.

This study makes use of 5-year crashes (2001 to 2005) that occurred on 89 urban freeway segments throughout the state of Florida in the United States. The primary objective is to help improve selection of countermeasures by identifying effective predictors of injury severity on urban freeways by applying the mixed logit approach to account for unobserved (or latent) effects. A second objective is to find out how the model's coefficients could vary from both the mixed logit model and the multinomial logit model. The latter model belongs to the same family of the mixed logit model and has been successfully applied in severity studies (Carson & Mannering, 2001). A third objective is to compare the goodness-of-fit from the fitted mixed and multinomial logit models to help in assessing the recommendation of a robust approach for severity analysis on urban freeways.

The next section of this paper provides a review of those studies that applied the mixed logit and multinomial logit approaches to model injury severity. This is followed by a description of the data and explored variables, and an overview of the mixed and multinomial logit models. The results from the fitted mixed and multinomial logit models are then presented and compared. The final section concludes with the key findings and provides recommendations for further research.

2. Literature review

This section reviews severity studies that applied the mixed logit model, the multinomial logit model, as well as severity analysis on freeways. One of the first studies that applied the mixed logit model for safety analysis is by Gkritza and Mannering (2008). They analyzed the safety belt usage in single- and multi-occupant vehicles using data collected from an observational roadside survey of safety belt use in Indiana. They hypothesized that the mixed logit approach offered more flexibility while capturing individual-specific heterogeneity from roadway characteristics, driver behavior, and vehicle types. The authors found that for single-occupant vehicles, male drivers, truck and van drivers, and those driving in the morning were less likely to use safety belts. For multiple-occupant vehicles, sport utility vehicle (SUV) drivers and drivers along an interstate highway were more likely that both front-seat occupants were restrained, while driving in a van and in the afternoon decreased this likelihood.

In another pioneering study, Milton et al. (2008) used the mixed logit model to study the injury severity distribution of crashes on highway segments in Washington. They found that volume-related variables such as average daily traffic, average daily truck traffic, truck percentage, number of interchanges per mile, and weather features such as snowfall, were better modeled as random parameters, while roadway variables, such as the number of horizontal curves, number of grade breaks per mile, and pavement friction, were better modeled as fixed. Afterwards, Pai et al. (2009) examined the characteristics of auto–motorcycle crashes related to gap-acceptance at T-junctions in the United Kingdom, using the British accident injury database from 1991 to 2005. The authors used the mixed logit approach to model the factors contributing to motorists' right-of-way (ROW) violation. It was concluded that motorcycles' ROW was more likely to be violated on non-built-up roads and in poor lighting conditions.

Several more recent studies can be found in the literature. Kim, Ulfarsson, Shankar, and Mannering (2010) used police-reported crashes from 1997 to 2000 from North Carolina to investigate pedestrian injury severity predictors using the mixed logit approach. They discovered several factors that more than double the pedestrian fatal injury risk, including darkness with no streetlights, trucks, freeway facilities, speeding involvement, and drinking while driving. Malyshkina and Mannering (2010) empirically investigated the effect of highway design exceptions on the frequency and severity of crashes. Design exceptions are defined as some cases for constructing or maintaining highways without meeting the specified guidelines (e.g., for design speed, lane width, shoulder width). Crash data were extracted from the State of Indiana's crash records for the period of 2003 to 2007. The authors classified the severity into three levels: fatal, injury (possible, evident, and disabling), and property damage only. They mapped these crashes on 35 segments that had design exceptions at bridges, and 13 segments that had design exceptions along the roadways.

Malyshkina and Mannering further used the mixed logit model and concluded that crashes in urban areas had a lower probability of injury crashes, whereas high speed limits were associated with higher injury crashes likelihood. Moreover, they found that the current process for granting design exceptions did not have a significant impact on crash frequency or severity. Anastasopoulos and Mannering (2011) compared the fixed and random parameter logit (or mixed logit) models in the context of their data structure using five-year data from interstate highways in the State of Indiana. The authors explored two types of data within each model: detailed crash-specific data and general non-detailed data, including the injury outcome of the crash and some roadway and traffic features. The authors concluded that the mixed logit model using less detailed data could provide an acceptable level of estimation accuracy.

Another recent study from Moore, Schneider, Savolainen, and Farzaneh (2011) looked into why the mixed logit model could help avoiding the shortcomings of previously explored severity models (e.g., the multinomial logit model). The authors explained that the main advantages of the mixed logit lie in its ability to relax the independence from irrelevant alternatives (IIA) property, as well as the heterogeneity in parameter estimates across the observations. The authors investigated both the multinomial logit and mixed logit models to identify those geometric, environmental, and crash type characteristics affecting bicyclists' injury severity at intersection and non-intersection locations. The data were extracted from 2002 to 2008 for bicycle-involved crashes in Ohio. For crashes occurring at intersection locations, the probability of bicyclist injury severity increases if the bicyclist is not wearing a helmet, if the motorist is under the influence of alcohol, if the motor vehicle involved in the crash is a van, if the motor vehicle strikes the side of the bicycle, and if the crash occurs on a horizontal curve that has a grade. For non-intersection location crashes, the likelihood of bicyclist severity increases if the bicyclist is under the influence of drugs, if the motorist is under the influence of alcohol, if the motor vehicle strikes the side of the bicycle, and if the motor vehicle involved in the crash is a truck.

The latest studies found in the literature have included Kim et al. (2011) and Ye and Lord (2011). Kim et al. (2011) developed a mixed logit model of driver injury severity in single-vehicle crashes to explore the effect of driver age on those particular crashes. They used crash data from 2003 to 2004 from the California Highway Patrol data records. The identified factors that increased fatal injury probability included older drivers (65 years or above), male drivers, drunk driving, older drivers driving older vehicles, and darkness with no streetlights. Ye and Lord (2011) compared the three most frequently-used severity models: the multinomial logit, the ordered probit, and the mixed logit models based on the required sample sizes for an effective estimation of the parameters. To achieve this objective, they used a Monte-Carlo approach using simulated and observed crash data. The authors concluded that the mixed logit model required the largest sample size. The recommended sample sizes for the multinomial logit, ordered probit, and mixed logit models are 1,000, 2,000, and 5,000, respectively.

From the studies that examined the multinomial logit model, Shankar and Mannering (1996) modeled the injury severity resulting from vehicle/motorcycle crashes using crash data from 1989 to 1994 in Washington. The authors concluded that the helmet usage could reduce injury severity in motorcycle crashes. Carson and Mannering (2001) analyzed the impact of ice-warning signs on crash severity along roadway sections and did not find them to have a clear impact. In another study, Ulfarsson and Mannering (2004) analyzed the