Accident Analysis and Prevention 85 (2015) 186–198

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

# Hierarchical Bayesian random intercept model-based cross-level interaction decomposition for truck driver injury severity investigations

Cong Chen[a], Guohui Zhang[a,*], Zong Tian[b], Susan M. Bogus[a], Yin Yang[c]

[a] Department of Civil Engineering, University of New Mexico, Albuquerque, NM 87131, USA
[b] Department of Civil and Environmental Engineering, University of Nevada, Reno, MS 0258, Reno, NV 89557, USA
[c] Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131, USA

ARTICLE INFO

*Article history:*
Received 23 March 2015
Received in revised form 5 September 2015
Accepted 9 September 2015
Available online 8 October 2015

*Keywords:*
Truck driver injury
Random intercept model
Unobserved heterogeneity
Cross-level interaction
Bayesian inference
Traffic safety

ABSTRACT

Traffic crashes occurring on rural roadways induce more severe injuries and fatalities than those in urban areas, especially when there are trucks involved. Truck drivers are found to suffer higher potential of crash injuries compared with other occupational labors. Besides, unobserved heterogeneity in crash data analysis is a critical issue that needs to be carefully addressed. In this study, a hierarchical Bayesian random intercept model decomposing cross-level interaction effects as unobserved heterogeneity is developed to examine the posterior probabilities of truck driver injuries in rural truck-involved crashes. The interaction effects contributing to truck driver injury outcomes are investigated based on two-year rural truck-involved crashes in New Mexico from 2010 to 2011. The analysis results indicate that the cross-level interaction effects play an important role in predicting truck driver injury severities, and the proposed model produces comparable performance with the traditional random intercept model and the mixed logit model even after penalization by high model complexity. It is revealed that factors including road grade, number of vehicles involved in a crash, maximum vehicle damage in a crash, vehicle actions, driver age, seatbelt use, and driver under alcohol or drug influence, as well as a portion of their cross-level interaction effects with other variables are significantly associated with truck driver incapacitating injuries and fatalities. These findings are helpful to understand the respective or joint impacts of these attributes on truck driver injury patterns in rural truck-involved crashes.



## 1. Introduction

Traffic crashes on roadways are one of the major threats to human safety and health, resulting in significant life and economic loss. In the U.S., there are 33,000 fatalities and 2.4 million injuries on average resulting from 5.6 million reported crashes annually (NHTSA, 2013a,b). Traffic crashes occurring on rural roadways are prone to inducing more severe injuries and fatalities, compared with those occurring in urban areas. 54% of total fatal crashes and 55% of total fatalities in the U.S. happen on rural roadways, where only 34% of vehicle miles traveled (VMT) occur (NHTSA, 2013a,b). Considerable research has been conducted to address rural traffic crashes from various perspectives regarding the features of roadways, vehicles, drivers, and natural environments, and corresponding laws and regulations have been legitimized to reduce rural crash risks (Cannon et al., 2009; Gallaher et al., 1989; Lord et al., 2005). More in-depth investigations revealed that 72% of fatal crashes in the U.S. happened on rural two-lane highways (NHTSA, 2013a,b), which has drawn increasing research attention. For instance, Cafiso et al. (2010) examined two-lane rural highway crash characteristics through synthetical models, taking into account the factors associated with safety performance such as exposure and context variables. Kashani and Mohaymany (2011) applied classification tree models to predict the injury severity patterns of two-lane rural roadway traffic accidents, and concluded that inadequate overtaking and seatbelt nonuse are the most significant contributors to severe crash injuries.

Due to the considerable sizes and weights of trucks, traffic crashes with truck(s) involved always cause more severe vehicle damage and injury outcomes. According to the Federal Motor Carrier Safety Administration (FMCSA) (FMCSA, 2014), truck crashes cost around $30 billion annually and 52% of truck-involved crashes result in driver incapacitating injuries or fatalities. Abundant studies were conducted to investigate the features of truck-involved

* Corresponding author.
*E-mail address:* guohui@unm.edu (G. Zhang).

http://dx.doi.org/10.1016/j.aap.2015.09.005
0001-4575/

crashes (Cantor et al., 2010; Chang and Chien, 2013; Chang and Mannering, 1999; Chen and Xie, 2014; Duke et al., 2010; Lee and Li, 2014; Lemp et al., 2011; Levine et al., 1999; Spielholz et al., 2008; Zaloshnja and Miller, 2004). For instance, Cantor et al. (2010) assessed the impacts of truck driver characteristics on crash severity outcomes and found that driver gender, age, weight, height, and employment stability are critical factors associated with crash severity. Lemp et al. (2011) discovered that an increasing number of trailers tend to increase the odds of severe injuries and fatalities. Although it is revealed that truck drivers themselves suffer less severe injuries than passenger vehicle drivers due to the protective effect from truck weight and size (Levine et al., 1999), truck drivers still endure higher potential of severe injury outcomes in traffic crashes than other workers (NIOSH, 2007). To address this issue, Shibuya et al. (2010) investigated the major factors related to non-traffic occupational injuries for truck drivers and established the hazard scenarios during loading and unloading procedures. According to FMCSA (2006), of all the crashes where the involved truck was assigned as the reason, 28% percent resulted from truck drivers' recognition issues and another 38% were induced by drivers' inappropriate decisions. Therefore, peer studies have been performed to address truck drivers' perception and behavior patterns toward traffic environments (Anderson et al., 1999; Rosenbloom et al., 2009). For instance, Rosenbloom et al. (2009) investigated the discrepancy in the approaches toward reckless driving for both truck drivers and non-truck drivers and the factors affecting truck drivers' reckless driving perception.

Even though crash injury severity analyses have been conducted for decades, there are still concerns and limitations that should raise researchers' attention. Mannering and Bhat (2014) summarized these research limitations in accident research, of which unobserved heterogeneity has been recognized as a critical issue in traffic safety research. Unobserved heterogeneity is defined as the unobservable factors or data that affect crash potential or severity, and they may generate biased estimations if their correlations with observed variables are not accounted for in model design (Mannering and Bhat, 2014). Random parameters models and finite-mixture (latent-class) models are two major types of models that are widely used in traffic safety analysis to address unobserved heterogeneity in crash data resulting from roadway features, driver demographic and behavior information, spatial and temporal variations, etc. (Anastasopoulos et al., 2012a,b; Chen and Tarko, 2014; Flask et al., 2014; Haleem and Gan, 2013; Islam and Mannering, 2006; Kim et al., 2013; Malyshkina and Mannering, 2010; Malyshkina et al., 2009; Xiong and Mannering, 2013). In this study, the cross-level interaction effects between crash-level and vehicle level variables are generally not observed in crash data, but may have significant influence on driver injury severity outcomes in traffic crashes. Random parameters models and finite mixture models are effective in examining the unobserved heterogeneity, but these interaction effects were generally neglected. Correlated random parameters models are able to model sophisticated interactions effects, but the cross-level interaction effects have not been systematically examined yet; the random intercept and slope models proposed by Snijders and Bosker (2000) provide robust designs to comprehensively examine these cross-level interactions effects. This cross-level interaction model configuration has been widely discussed in statistics (Gökalp and Kiliç, 2013; Iyit and Genc, 2011; Wagner and Duller, 2012) and applied in clinical data analysis (Iyit and Genc, 2011), but has never been utilized in traffic safety analyses. Therefore, this study is conducted as complementary research to test the applicability of an alternative method to investigate unobserved heterogeneity.

In traffic safety analyses, the traditional maximum likelihood estimation (MLE) method quantifies parameters of interest directly from the studied dataset and ignores the structured disturbance from the hierarchical crash data structure (Huang and Abdel-Aty, 2010), which compromises the precision and transferability of the estimated results. Therefore, based on the hierarchical nature of traffic crash data, Gelman and Hill (2007) defined hierarchical regression modeling for traffic safety regressions, where parameters of interests were given probability models. The Bayesian inference method generates parameters of interest with probability distributions based on prior information and the studied dataset, which outperforms the traditional MLE method. With the prevalence of hierarchical model applications in traffic safety analyses, the Bayesian inference method has a high level of implementation in traffic safety analyses (Ahmed et al., 2011; Haque et al., 2010; Huang et al., 2014, 2011; Ma and Kockelman, 2006; MacNab, 2004, 2003; Yu and Abdel-Aty, 2014, 2013). Hence, the parameters representing these individual main effects and interaction effects in this study were assumed as random parameters following certain pre-defined distributions and estimated through Bayesian inference method with a Monte Carlo Markov Chain (MCMC) simulation algorithm. The above literature provides a comprehensive understanding regarding rural crashes, truck-involved crashes, truck driver characteristics, unobserved heterogeneity in crash data, and hierarchical Bayesian modeling. Given the severe injury outcomes in rural and truck crashes, this article aims to examine the injury severity patterns of truck drivers in rural traffic crashes. Three injury severity levels were considered in this study: no injury (property damage only) (coded as **N**), non-incapacitating injury (coded as **I**), and incapacitating injury or fatality (coded as **F**). A modified hierarchical Bayesian random intercept model based on Snijders and Bosker (2000) is utilized to estimate the influence of heterogeneous crash-level and vehicle/driver-level factors contributable to truck driver injury severities, and to explore the unobserved heterogeneity in terms of cross-level interactions between different data hierarchies. The rest of this article is organized as follows: Section 2 presents a comprehensive literature review regarding unobserved heterogeneity analysis in traffic safety research. Section 3 provides detailed information of the studied dataset, followed by Section 4 describing the methodology design and model specifications. The modeling results are explicitly discussed in Section 5, accompanied by Section 6 summarizing the limitations of this study. This research is concluded in Section 7.

## 2. Unobserved heterogeneity in crash data analysis

Unobserved heterogeneity has been recognized as a critical issue in traffic safety research. Unobserved heterogeneity is defined as the unobservable factors or data that affect crash potential or severity, and they may generate biased estimations if their correlations with observed variables are not accounted for in model design (Mannering and Bhat, 2014). The unobserved heterogeneity could be attributed from different types of factors, including roadways (Flask et al., 2014; Haleem and Gan, 2013; Malyshkina and Mannering, 2010; Morgan and Mannering, 2011), drivers' demographic and behavior characteristics (Haleem and Gan, 2013; Islam and Mannering, 2006; Kim et al., 2013, 2010; Morgan and Mannering, 2011; Ulfarsson and Mannering, 2004), spatial and temporal variations (Malyshkina and Mannering, 2009; Malyshkina et al., 2009; Ukkusuri et al., 2011; Xiong et al., 2014; Xu and Huang, 2015), etc. For instance, Kim et al. (2010) evaluate pedestrian injury severity patterns in pedestrian-vehicle crashes considering the unobserved pedestrian heterogeneity regarding health, strength and behavior. Anastasopoulos et al. (2012a,b) investigated traffic accident rate patterns accounting for the unobserved heterogeneity effects of highway segments. Xiong et al. (2014) examined crash injury severity patterns based on the heterogeneous temporal influence of roadway segment features.