Analytic Methods in Accident Research 14 (2017) 22–40

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

A Modified Rank Ordered Logit model to analyze injury severity of occupants in multivehicle crashes

Shelley Bogue, Rajesh Paleti *, Lacramioara Balan

Department of Civil & Environmental Engineering, Old Dominion University, 135 Kaufman Hall, Norfolk, VA 23529, United States

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 28 February 2017
Accepted 1 March 2017
Available online 10 March 2017

Keywords:
Rank ordered logit
Mixed logit
Injury severity
Multiple occupants
Parameter heterogeneity

ABSTRACT

The current study developed a simultaneous model of injury severity outcomes of all occupants in multi-vehicle crashes including all the drivers and the passengers of all vehicles involved in a crash. Specifically, a Modified Rank Ordered Logit (MROL) methodology that can predict the relative order of occupant injury severity as well as the actual injury severity was developed. The final model captures the effects of several key occupant, vehicle, and accident level variables on four possible levels of injury severity. The results indicate the presence of accident-specific unobserved factors that influence the severity outcomes of all people involved in the crash as well as unobserved heterogeneity in the effect of key covariates including occupant's gender and speed limit. The performance of the MROL model was compared with the traditional mixed multinomial logit (MMNL) model that is the most commonly used model for injury severity analysis. Overall, the results demonstrate superior predictive ability of the MROL model in comparison to the MMNL model. The traditional MMNL model performed satisfactory in terms of replicating the simple shares of different injury severity levels across all occupants. However, the performance of the MMNL model dropped significantly when the observed and predicted shares were compared for combinations of injury severity levels among crashes involving multiple occupants. Lastly, elasticity effects were computed to demonstrate considerably different policy implications of the MROL and MMNL models.

1. Introduction

Past research has been highly geared towards modeling injury severity of the driver or the most severely injured person in a crash (Chen and Chen, 2011; Lee and Li, 2014; Kim et al., 2013; Anastasopoulos and Mannering, 2011; Paleti et al., 2010; Xie et al., 2012). Even studies that attempted to model injury severity of multiple occupants restricted themselves to drivers or occupants of single and two vehicle crashes primarily due to the methodological complications associated with joint modeling of severity outcomes of multiple occupants. In some cases, information about other occupants' injury-severity outcomes, or single *versus* multiple crash occurrence was not available. In such cases, data availability was the reason for restricting the analysis to crash or driver injury severity. However, in recent times, crash databases have started to collect detailed information of all occupants involved in a crash enabling multi-occupant analysis. (Savolainen et al., 2011) provides a detailed review of different statistical techniques that were used to analyze highway crash injury severities. One of the critical methodological aspects identified by this study is within-crash correlation among different crash-injury observations.

* Corresponding author.
E-mail addresses: sbogu001@odu.edu (S. Bogue), rpaleti@odu.edu (R. Paleti), lbala001@odu.edu (L. Balan).

http://dx.doi.org/10.1016/j.amar.2017.03.001

Ignoring crash-level unobserved factors that generate this within-crash correlation during model estimation can result in biased parameter estimates. So, understandably, recent studies that analyzed severity of multiple occupants used simultaneous modeling methods to account for within crash correlation. For instance, (Rana et al., 2010) created a copula-based model that ties together the risk propensities of drivers in two vehicle crashes. Similarly, (Abay et al., 2013) developed mixed logit models for analyzing severity outcomes of drivers in two vehicle crashes while accounting for seat belt endogeneity. Along similar lines, (Zhu and Srinivasan, 2011) developed a panel heteroskedastic ordered model to analyze injury severity of multiple occupants involved in large truck crashes. One study of direct relevance to the current analysis is (Eluru et al., 2010) that modeled injury severity of multiple occupants in a vehicle based on their seating positions including driver, front-seat passenger, rear-center passenger, and other rear seat configuration. However, while this study accounts for within-vehicle correlation, within-crash correlation across occupants in multiple vehicles of the same crash is ignored. Moreover, this study focused only on crashes involving one or two vehicles.

A common feature of these studies that attempted to model severity outcomes of multiple occupants is an econometric framework that has separate propensity/severity functions for modeling the severity level of each occupant where the unobserved components of these propensity/severity functions are tied together using a joint distribution. In these models, injury severity sustained by one occupant in a crash does not directly influence the risk propensity of other occupants. Any dependency or correlation arises purely out of the correlation in the unobserved components of underlying propensity or severity functions. While these models can predict the profile of severity levels by each occupant, they cannot replicate the complete relative profile of severity levels sustained by occupants in the same crash. That is to say, if we consider single vehicle crashes, the predicted profiles of severity levels sustained by passengers and drivers will be reasonably close to the respective observed profiles; however, if we look at the percentage of crashes in which the driver is more severely injured than the passenger (or *vice versa*) and compare the predicted and observed profiles, there will be significant deviations.

One way to introduce conditionality is to assume a pre-determined order in which severity levels of occupants are modeled. For instance, model severity levels of passengers conditional on drivers. However, the number of assumptions increase considerably as the number of occupants and the number of vehicles involved in the crash increase. Also, it is difficult to justify why one particular sequence or order is better than other sequences of occupants. However, we also cannot deny that, in any crash, there are occupants who are at a higher risk of sustaining severe injuries compared to other occupants. So, there is an intrinsic ordering in the injury severity levels sustained by occupants in multi-vehicle crashes. Furthermore, this intrinsic ordering of severity levels sustained by occupants can vary across crashes. So, instead of assuming a pre-determined sequence, we want the data to inform us about the relative ordering of severity levels sustained by occupants involved in the crash. While doing so, the model must also account for potential crash-specific common unobserved factors that influence injury severity of occupants in the same crash. These unobserved factors can also moderate the influence of other factors that affect severity outcome of all occupants involved in the same crash.

In this context, the current paper aims to bridge the existing gap in the safety literature on joint modeling of severity outcomes of all occupants in multi-vehicle crashes by accomplishing the following five study objectives. *First*, develop a Modified Rank Ordered Logit (MROL) model that is capable of ranking all occupants involved in a crash in the decreasing order of risk propensity as well as predicting their injury severity outcomes. The MROL framework is easily scalable for multi-occupant and multi-vehicle crashes unlike earlier methods that were primarily geared towards single and two vehicle crashes. *Second*, estimate the parameters of the MROL model using the national crash database, the 2013 National Automotive Sampling System (NASS) General Estimates System (GES) of the National Highway Traffic Safety Administration (NHTSA). The 2013 GES data includes detailed information including occupant, vehicle, and accident characteristics for a total of 34,782 crashes, 62,879 vehicles, and 86,519 occupants. *Third*, explore within crash correlation and potential heterogeneity in the parameter estimates of the MROL model due to crash-level unobserved factors that moderate the influence of difference exogenous variables on the injury severity outcomes of all people involved in the same crash. *Fourth*, estimate the Mixed Multinomial Logit (MMNL) model, the most commonly used model in the safety literature for injury severity modeling and demonstrate the superior predictive performance of the MROL model by comparing the forecasts from the MROL and MMNL models. *And lastly*, compare the policy implications of the MROL and MMNL models by computing elasticity effects.

2. Methodology

Past literature suggests that researchers have primarily used two methodological frameworks for analyzing injury severity, namely unordered and ordered. In the unordered response framework, each possible severity level is assumed to be associated with a latent severity function and the observed severity outcome is the severity level with the maximum severity function value. The Generalized Extreme Value (GEV) models that include the standard multinomial logit (MNL), nested logit (NL) and mixed logit (ML) models belong to the unordered response modeling framework (Wu et al., 2014b; Kim et al., 2013; Milton et al., 2008; Huang et al., 2016; Behnood and Mannering, 2015). In the ordered response framework, a single latent risk propensity function is mapped into the observed severity outcome by threshold parameters. The standard ordered response (OR) model and its generalizations including the partial proportional odds (PPO) and the generalized ordered response (GOR) models belong to the class of ordered response models (Mooradian et al., 2013; Wang and Abdel-Aty, 2008; Yasmin et al., 2014a; Eluru et al., 2008; Chen et al., 2016a; Russo et al., 2014). Recent comparisons of the ordered and the unordered response models for severity modeling found that these two methods are equivalent in terms of the